

SEQUENCE LISTING

- (1) GENERAL INFORMATION:
- (i) APPLICANT: Iris Pecker, Israel Vlodavsky and Elena Feinstein
 - (ii) TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A POLYPEPTIDE HAVING HEPARANASE ACTIVITY AND EXPRESSION OF SAME IN GENETICALLY MODIFIED CELLS
 - (iii) NUMBER OF SEQUENCES: 47
 - (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Mark M. Friedman c/o Anthony Castorina
 - (B) STREET: 2001 Jefferson Davis Highway, Suite 207
 - (C) CITY: Arlington
 - (D) STATE: Virginia
 - (E) COUNTRY: United States of America
 - (F) ZIP: 22202
 - (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk
 - (B) COMPUTER: Twinhead* Slimnote-890TX
 - (C) OPERATING SYSTEM: MS DOS version 6.2, Windows version 3.11
 - (D) SOFTWARE: Word for Windows version 2.0 converted to an ASCII file
 - (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE:
 - (C) CLASSIFICATION:
 - (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: 08/922,170
 - (B) FILING DATE: 2 SEP 1997
 - (A) APPLICATION NUMBER: 09/109,386
 - (B) FILING DATE: 10 JUL 1998
 - (A) APPLICATION NUMBER: PCT/US98/17954
 - (B) FILING DATE: 31 AUG 1998
 - (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Friedman, Mark M.
 - (B) REGISTRATION NUMBER: 33,883
 - (C) REFERENCE/DOCKET NUMBER: 910/14
 - (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: 972-3-5625553
 - (B) TELEFAX: 972-3-5625554
 - (C) TELEX:
- (2) INFORMATION FOR SEQ ID NO:1:
- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 27
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
CCATCTCAAT ACGACTCACT ATAGGGC 27
- (2) INFORMATION FOR SEQ ID NO:2:
- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 24
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GTAGTGATGC CATGTAAGTG AATC 24

- ```
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:
 ACTCACTATA GGGCTCGAGC GGC 23
```

## (2) INFORMATION FOR SEQ ID NO:9:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1721  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CTAGAGCTTT CGACTCTCCG CTCGCGGCA GCTGGCGGG GAGCAGCCCA GGTGAGCCCA 60  
 AGATGGTGCT GCGCTCGAAG CTTGGCTGCG CGCGCGCGCT GATGCTGCTG CTCTGGGGG 120  
 CGCTGGCTCC CCTCTCCCTT GCGCGCTCG CCCGACCTGC GCAAGCAGAG GACGTCTGG 180  
 ACCTGGACTT CTTCAACCCG GAGCGCGCTG ACCTGGTGAG CCGCTCGTTC CTGTCCGTC 240  
 CCATTGACGC CACCTGGGCC ACGGACCCGC GGTTCCTCAT CCTCCTGGGT TCTCCAAAGC 300  
 TTGTAACCTT GGCAGAGGC TTGCTCCTG CGTACCTGAG GTTTGGTGCG ACCAAGACAG 360  
 ACTTCTTAAT TTTCGATCCC AAGAAGGAAT CAACCTTTGA AGAGAGAAGT TACTGGCAAT 420  
 CTCAGTCAA CCAGGATATT TGCAATATG GATCCATCCC TCCTGATGTG GAGGAGAAGT 480  
 TAGCGTTGGA ATGGCCCTAC CAGGAGCAAT TGCTACTCGG AGAACACTAC CAGAAAAAGT 540  
 TCAAGAACAG CACCTACTCA AGAAGCTCTG TAGATCTGCT ATACACTTTT GCAAACTGCT 600  
 CAGGACTGGA CTTGATCTTT GGCCTAANTG CGTTATTAAG AACAGCAGAT TTGCAGTGGA 660  
 ACAGTTCTAA TGCTCAGTTG CTCCTGGACT ACTGCTTTC CAAGGGGTAT AACATTTCCT 720  
 GGGAACTAGG CAATGAACCT AACAGTTTCC TTAAGAAGCG TGATATTTC ATCAATGGGT 780  
 CGCAGTTAGG AGAAGATTAT ATTCAATTGC ATAACTTCT AAGAAGTGCC ACCTTCAAAA 840  
 ATGCAAAACT CTATGGTCTT GATGTTGGTC AGCCTCGAAG AAAGACGGCT AAGATGCTGA 900  
 AGAGCTTCTT GAAGGCTGGT GGAGAAGTGA TTGATTCAAT TACATGGCAT CACTACTATT 960  
 TGAATGGAGG GACTGCTACC AGGGAAGATT TTCTAAACC TGATGTATTG GACATTTTTC 1020  
 TTTCACTCTG GCACAAAGTT TTCCAGGTGG TTGAGAGCAC CAGGCTGGCC AAGAAGTCT 1080  
 GGTGAGAGA AACAGCTCT GCATATGGAG GCGGAGCGCC CTGTCTATCC GACACCTTTG 1140  
 GACTGGTCTT TATGTGGCTG GATAAATGG GCCTGTGACG CCGAATGGGA ATAGAAGTGG 1200  
 TGATGAGGCA AGTATTCTTT GGAGCAGGAA ACTACCTATT AGTGGATGAA AACTTCGATC 1260  
 CTTTACCTGA TTATTGGCTA TCTCTTCTGT TCAAGAAATT GGTGGGCACC AAGGTGTTAA 1320  
 TGCAAGCGT GCAAGGTTCA AAGAAGAAGA AGCTTCGAGT ATACCTTCAT TGCACAAAC 1380  
 CTGCAATACC AAGGTATAAA GAAGGAGATT TAACCTCTGA TGCCATAAAC CTCCTAACG 1440  
 TCACCAAGTA CTTGCGGTTA CCTATCCTT TTTCTAAAC GCAAGTGGAT AAATACCTTC 1500  
 TAAGACCTTT GGGACCTCAT GGATTACTTT CCAAACTGT CCAACTCAAT GGTCTAACTC 1560  
 TAAAGATGGT GGATGATCAA ACCTTGCCAC CTTTAAATGGA AAAAAGCTCT CGGCCAGGAA 1620  
 GTTCACTGGG CTGCGCAGCT TTCTCATATA GTTTTTTGT GATAAGAAAT GCCAAAGTGT 1680  
 CTGCTTGACT CTGAAAAATA AATATACTAG TCCTGACACT G 1721

## (2) INFORMATION FOR SEQ ID NO:10:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 543  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met Leu Leu Arg Ser Lys Pro Ala Leu Pro Pro Pro Leu Met Leu Leu  
 5 10 15  
 Leu Leu Gly Pro Leu Gly Pro Leu Ser Pro Gly Ala Leu Pro Arg Pro  
 20 25 30  
 Ala Gln Ala Gln Asp Val Val Asp Leu Asp Phe Phe Thr Gln Glu Pro  
 35 40 45  
 Leu His Leu Val Ser Pro Ser Phe Leu Ser Val Thr Ile Asp Ala Asn  
 50 55 60  
 Leu Ala Thr Asp Pro Arg Phe Leu Ile Leu Leu Gly Ser Pro Lys Leu

65                      70                      75                      80  
 Arg Thr Leu Ala Arg Gly Leu Ser Pro Ala Tyr Leu Arg Phe Gly Gly  
                             85                      90                      95  
 Thr Lys Thr Asp Phe Leu Ile Phe Asp Pro Lys Lys Glu Ser Thr Phe  
                             100                      105                      110  
 Glu Glu Arg Ser Tyr Trp Gln Ser Gln Val Asn Gln Asp Ile Cys Lys  
                             115                      120                      125  
 Tyr Gly Ser Ile Pro Pro Asp Val Glu Glu Lys Leu Arg Leu Glu Trp  
                             130                      135                      140  
 Pro Tyr Gln Glu Gln Leu Leu Leu Arg Glu His Tyr Gln Lys Lys Phe  
                             145                      150                      155                      160  
 Lys Asn Ser Thr Tyr Ser Arg Ser Ser Val Asp Val Leu Tyr Thr Phe  
                             165                      170                      175  
 Ala Asn Cys Ser Gly Leu Asp Leu Ile Phe Gly Leu Asn Ala Leu Leu  
                             180                      185                      190  
 Arg Thr Ala Asp Leu Gln Trp Asn Ser Ser Asn Ala Gln Leu Leu Leu  
                             195                      200                      205  
 Asp Tyr Cys Ser Ser Lys Gly Tyr Asn Ile Ser Trp Glu Leu Gly Asn  
                             210                      215                      220  
 Glu Pro Asn Ser Phe Leu Lys Lys Ala Asp Ile Phe Ile Asn Gly Ser  
                             225                      230                      235                      240  
 Gln Leu Gly Glu Asp Tyr Ile Gln Leu His Lys Leu Leu Arg Lys Ser  
                             245                      250                      255  
 Thr Phe Lys Asn Ala Lys Leu Tyr Gly Pro Asp Val Gly Gln Pro Arg  
                             260                      265                      270  
 Arg Lys Thr Ala Lys Met Leu Lys Ser Phe Leu Lys Ala Gly Gly Glu  
                             275                      280                      285  
 Val Ile Asp Ser Val Thr Trp His His Tyr Tyr Leu Asn Gly Arg Thr  
                             290                      295                      300  
 Ala Thr Arg Glu Asp Phe Leu Asn Pro Asp Val Leu Asp Ile Phe Ile  
                             305                      310                      315                      320  
 Ser Ser Val Gln Lys Val Phe Gln Val Val Glu Ser Thr Arg Pro Gly  
                             325                      330                      335  
 Lys Lys Val Trp Leu Gly Glu Thr Ser Ser Ala Tyr Gly Gly Gly Ala  
                             340                      345                      350  
 Pro Leu Leu Ser Asp Thr Phe Ala Ala Gly Phe Met Trp Leu Asp Lys  
                             355                      360                      365  
 Leu Gly Leu Ser Ala Arg Met Gly Ile Glu Val Val Met Arg Gln Val

370 375 380

Phe Gly Ala Gly Asn Tyr His Leu Val Asp Glu Asn Phe Asp Pro  
385 390 395 400

Leu Pro Asp Tyr Trp Leu Ser Leu Leu Phe Lys Lys Leu Val Gly Thr  
405 410 415

Lys Val Leu Met Ala Ser Val Gln Gly Ser Lys Arg Arg Lys Leu Arg  
420 425 430

Val Tyr Leu His Cys Thr Asn Thr Asp Asn Pro Arg Tyr Lys Glu Gly  
435 440 445

Asp Leu Thr Leu Tyr Ala Ile Asn Leu His Asn Val Thr Lys Tyr Leu  
450 455 460

Arg Leu Pro Tyr Pro Phe Ser Asn Lys Gln Val Asp Lys Tyr Leu Leu  
465 470 475 480

Arg Pro Leu Gly Pro His Gly Leu Leu Ser Lys Ser Val Gln Leu Asn  
485 490 495

Gly Leu Thr Leu Lys Met Val Asp Asp Gln Thr Leu Pro Pro Leu Met  
500 505 510

Glu Lys Pro Leu Arg Pro Gly Ser Ser Leu Gly Leu Pro Ala Phe Ser  
515 520 525

Tyr Ser Phe Phe Val Ile Arg Asn Ala Lys Val Ala Cys Ile  
530 535 540 543

(2) INFORMATION FOR SEQ ID NO:11:

(1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1721

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

CT AGA GCT TTC GAC 14

TCT CCG CTG CGC GGC AGC TGG CGG GGG GAG CAG CCA GGT GAG CCC AAG 62

ATG CTG CTG CGC TCG AAG CCT GCG CTG CCG CCG CCG CTG ATG CTG CTG 110  
Met Leu Leu Arg Ser Lys Pro Ala Leu Pro Pro Pro Leu Met Leu Leu  
5 10 15

CTC CTG GGG CCG CTG GGT CCC CTC TCC CCT GGC GCC CTG CCC CGA CCT 158  
Leu Leu Gly Pro Leu Gly Pro Leu Ser Pro Gly Ala Leu Pro Arg Pro  
20 25 30

CGC CAA GCA CAG GAC GTC GTG GAC CTG GAC TTC TTC ACC CAG GAG CCG 206  
Ala Gln Ala Gln Asp Val Val Asp Leu Asp Phe Phe Thr Gln Glu Pro  
35 40 45

CTG CAC CTG GTG AGC CCC TCG TTC CTG TCC GTC ACC ATT GAC GCC AAC 254

Leu His Leu Val Ser Pro Ser Phe Leu Ser Val Thr Ile Asp Ala Asn 302  
 50 55 60  
 CTG GCC ACG GAC CCG CGG TTC CTC ATC CTC CTG GGT TCT CCA AAG CTT 302  
 Leu Ala Thr Asp Pro Arg Phe Leu Ile Leu Leu Gly Ser Pro Lys Leu  
 65 70 75 80  
 CGT ACC TTG GCC AGA GGC TTG TCT CCT GCG TAC CTG AGG TTT GGT GGC 350  
 Arg Thr Leu Ala Arg Gly Leu Ser Pro Ala Tyr Leu Arg Phe Gly Gly  
 85 90 95  
 ACC AAG ACA GAC TTC CTA ATT TTC GAT CCC AAG AAG GAA TCA ACC TTT 398  
 Thr Lys Thr Asp Phe Leu Ile Phe Asp Pro Lys Lys Glu Ser Thr Phe  
 100 105 110  
 GAA GAG AGA AGT TAC TGG CAA TCT CAA GTC AAC CAG GAT ATT TGC AAA 446  
 Glu Glu Arg Ser Tyr Trp Gln Ser Gln Val Asn Gln Asp Ile Cys Lys  
 115 120 125  
 TAT GGA TCC ATC CCT CCT GAT GTG GAG GAG AAG TTA CGG TTG GAA TGG 494  
 Tyr Gly Ser Ile Pro Pro Asp Val Glu Glu Lys Leu Arg Leu Glu Trp  
 130 135 140  
 CCC TAC CAG GAG CAA TTG CTA CTC CGA GAA CAC TAC CAG AAA AAG TTC 542  
 Pro Tyr Gln Glu Gln Leu Leu Arg Glu His Tyr Gln Lys Lys Phe  
 145 150 155 160  
 AAG AAC AGC ACC TAC TCA AGA AGC TCT GTA GAT GTG CTA TAC ACT TTT 590  
 Lys Asn Ser Thr Tyr Ser Arg Ser Ser Val Asp Val Leu Tyr Thr Phe  
 165 170 175  
 GCA AAC TGC TCA GGA CTG GAC TTG ATC TTT GGC CTA AAT GCG TTA TTA 638  
 Ala Asn Cys Ser Gly Leu Asp Leu Ile Phe Gly Leu Asn Ala Leu Leu  
 180 185 190  
 AGA ACA GCA GAT TTG CAG TGG AAC AGT TCT AAT GCT CAG TTG CTC CTG 686  
 Arg Thr Ala Asp Leu Gln Trp Asn Ser Ser Asn Ala Gln Leu Leu Leu  
 195 200 205  
 GAC TAC TGC TCT TCC AAG GGG TAT AAC ATT TCT TGG GAA CTA GGC AAT 734  
 Asp Tyr Cys Ser Ser Lys Gly Tyr Asn Ile Ser Trp Glu Leu Gly Asn  
 210 215 220  
 GAA CCT AAC AGT TTC CTT AAG AAG GCT GAT ATT TTC ATC AAT GGG TCG 782  
 Glu Pro Asn Ser Phe Leu Lys Lys Ala Asp Ile Phe Ile Asn Gly Ser  
 225 230 235 240  
 CAG TTA GGA GAA GAT TAT ATT CAA TTG CAT AAA CTT CTA AGA AAG TCC 830  
 Gln Leu Gly Glu Asp Tyr Ile Gln Leu His Lys Leu Leu Arg Lys Ser  
 245 250 255  
 ACC TTC AAA AAT GCA AAA CTC TAT GGT CCT GAT GTT GGT CAG CCT CGA 878  
 Thr Phe Lys Asn Ala Lys Leu Tyr Gly Pro Asp Val Gly Gln Pro Arg  
 260 265 270  
 AGA AAG ACG GCT AAG ATG CTG AAG AGC TTC CTG AAG GCT GGT GGA GAA 926  
 Arg Lys Thr Ala Lys Met Leu Lys Ser Phe Leu Lys Ala Gly Gly Glu

275

280

285

GTG ATT GAT TCA GIT ACA TGG CAT CAC TAC TAT TTG AAT GGA CGG ACT 974  
 Val Ile Asp Ser Val Thr Trp His His Tyr Tyr Leu Asn Gly Arg Thr  
 290 295 300

GCT ACC AGG GAA GAT TTT CTA AAC OCT GAT GTA TTG GAC ATT TTT ATT 1022  
 Ala Thr Arg Glu Asp Phe Leu Asn Pro Asp Val Leu Asp Ile Phe Ile  
 305 310 315 320

TCA TCT GTG CAA AAA GTT TTC CAG GTG GTT GAG AGC ACC AGG CCT GGC 1070  
 Ser Ser Val Gln Lys Val Phe Gln Val Val Glu Ser Thr Arg Pro Gly  
 325 330 335

AAG AAG GTC TGG TTA GGA GAA ACA AGC TCT GCA TAT GGA GGC GGA GCG 1118  
 Lys Lys Val Trp Leu Gly Glu Thr Ser Ser Ala Tyr Gly Gly Ala  
 340 345 350

CCC TTG CTA TCC GAC ACC TTT GCA GCT GGC TTT ATG TGG CTG GAT AAA 1166  
 Pro Leu Leu Ser Asp Thr Phe Ala Ala Gly Phe Met Trp Leu Asp Lys  
 355 360 365

TTG GGC CTG TCA GCC CGA ATG GGA ATA GAA GTG GTG ATG AGG CAA GTA 1214  
 Leu Gly Leu Ser Ala Arg Met Gly Ile Glu Val Val Met Arg Gln Val  
 370 375 380

TTC TTT GGA GCA GGA AAC TAC CAT TTA GTG GAT GAA AAC TTC GAT CCT 1262  
 Phe Phe Gly Ala Gly Asn Tyr His Leu Val Asp Glu Asn Phe Asp Pro  
 385 390 395 400

TTA CCT GAT TAT TGG CTA TCT CTT CTG TTC AAG AAA TTG GTG GGC ACC 1310  
 Leu Pro Asp Tyr Trp Leu Ser Leu Leu Phe Lys Lys Leu Val Gly Thr  
 405 410 415

AAG GTG TTA ATG GCA AGC GTG CAA GGT TCA AAG AGA AGG AAG CTT CGA 1358  
 Lys Val Leu Met Ala Ser Val Gln Gly Ser Lys Arg Arg Lys Leu Arg  
 420 425 430

GTA TAC CTT CAT TGC ACA AAC ACT GAC AAT CCA AGG TAT AAA GAA GGA 1406  
 Val Tyr Leu His Cys Thr Asn Thr Asp Asn Pro Arg Tyr Lys Glu Gly  
 435 440 445

GAT TTA ACT CTG TAT GCC ATA AAC CTC CAT AAC GTC ACC AAG TAC TTG 1454  
 Asp Leu Thr Leu Tyr Ala Ile Asn Leu His Asn Val Thr Lys Tyr Leu  
 450 455 460

CGG TTA CCC TAT CCT TTT TCT AAC AAG CAA GTG GAT AAA TAC CTT CTA 1502  
 Arg Leu Pro Tyr Pro Phe Ser Asn Lys Gln Val Asp Lys Tyr Leu Leu  
 465 470 475 480

AGA CCT TTG GGA CCT CAT GGA TTA CTT TCC AAA TCT GTC CAA CTC AAT 1550  
 Arg Pro Leu Gly Pro His Gly Leu Leu Ser Lys Ser Val Gln Leu Asn  
 485 490 495

GGT TCA ACT CTA AAG ATG GTG GAT GAT CAA ACC TTG CCA CCT TTA ATG 1598  
 Gly Leu Thr Leu Lys Met Val Asp Asp Gln Thr Leu Pro Pro Leu Met  
 500 505 510

GAA AAA CCT CTC CGG CCA GGA AGT TCA CTG GGC TTG CCA GCT TTC TCA 1646  
 Glu Lys Pro Leu Arg Pro Gly Ser Ser Leu Gly Leu Pro Ala Phe Ser  
 515 520 525

TAT AGT TTT TTT GTG ATA AGA AAT GCC AAA GTT GCT GCT TGC ATC TGA 1694  
 Tyr Ser Phe Phe Val Ile Arg Asn Ala Lys Val Ala Ala Cys Ile  
 530 535 540 543

AAA TAA AAT ATA CTA GTC CTG ACA CTG 1721

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 824
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12

CTGGCAAGAA GGCTGGGTG GGAGAGACGA GCTCAGCTTA CGGTGGCGGT GCACCCCTGC 60  
 TGTCACAACAC CTTTGCAGCT GGCTTTATGT GGCTGGATAA ATTGGCCCTG TCAGCCCCAGA 120  
 TGGGCATAGA AGTCGTGATG AGGCAGGTGT TCTTCGGAGC AGGCAACTAC CACTTAGTGG 180  
 ATGAAAACCTT TGAGCCTTTA CCTGATTACT GGCTCTCTCT TCTGTTCAGG AAACCTGGTAG 240  
 GTCCAGGGT GTTACTGTCA AGAGTGAAAG GCCCAGACAG GAGCAAACTC CGAGTGTATC 300  
 TCACTGCAC TAACTCTAT CACCCACGAT ATCAGGAAGG AGATCTAATC CTGTATGTCC 360  
 TGAACCTCCA TAATGTCAAC AAGCACTTGA AGGTACCGCC TCCGTGTTC AGGAACACAG 420  
 TGAATACGTA CCTTCTGAAG CCTTCGGGGC CGGATGGATT ACTTCCAAA TCTGTCCAA 480  
 TGAACGGTCA AATTCTGAAG ATGGTGGATG AGCAGACCTT GCCAGCTTTG ACAGAAAAC 540  
 CTCTCCCGGC AGGAAGTGCA CTAAGCCTGC CTGCCTTTTC CTATGGTTT TTTGTCAAA 600  
 GAAATGCCAA AATCGTGCT TGTATATGAA AATAAAAGGC ATACGGTACC CCTGAGACAA 660  
 AAGCCGAGGG GGGTGTATT CATAAAACAA AACCTAGTT TAGGAGGCCA CCTCCTTGGC 720  
 GAGTTCAGCA GCTTCGGGAG GGTGGGGTAC ACTTCAGTAT TACATTCAGT GTGGGTCTCT 780  
 CTCTAAGAGG AATACCTGAG GTGGTGACAG TTAATAGCAC TGTG 824

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1899
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13

GGGAAACGA GCAAGGAAGT AGGAGAGAGC CGGCAGGCG GGGCGGGGT GGGATGGAG 60  
 CAGTGGGAGG GATGAGAGAG AGGAGTGGGA GGGATGGAG GGCAGTGGG AGGGGTGAG 120  
 AGCGGTAAAG GGGCGAGGA AAGAGAGAAA GGGCGTGGG GCTCGCGGG AGGAAGTGCT 180  
 AGAGCTCTCG ACTCTCCGCT CGCGCGCAGC TGGCGGGGG AGCAGCGAGG TGAGCCCAAG 240  
 ATGCTGCTGC GCTCGAAGCC TGCCTGCCC CGCGCGCTGA TGCTGCTGCT CTTGGGGCCG 300  
 AGCGTTCGCC TCTCCCTGG CGCCCTGCC CGACTGTGGC AAGCACAGGA CGTGTGGAG 360  
 CTGCACTTCT TCACCCAGGA GCGCTGCAC CTGGTGAGCC CCTCGTTCT GTCCGTCAAC 420  
 ATTGACGCA AACTGGCCAC AGACCCGGG TTCTCATCC TCTGGGTTT TCCAAAGCTT 480  
 CGTACCTTGG CAGAGGCTT GTCTCTGCG TACCTGAGST TTGGTGGCAG CAAGACAGAG 540  
 TTCTATATT TCGATCCCAA GAAGCAATCA ACCTTTGAAG AGAGAAGTTA CTGGCAATCT 600  
 CAAGTCAACC AGGATATTTG CAAATATGGA TCCATCCCTC CTGATGTGGA GGAGAAGTTA 660  
 CGGTGGGAAT GGCCCTACCA GGAGCAATTG CTACTCCGAG AACACTACCA GAAAGAGTTC 720  
 AAGAACAGCA CTACTACAG AGCTCTGTA GATGTGCTAT ACACTTTTCG AAACGTGCTCA 780  
 GGACTGGAAT TGATCTTGG CTAATAGTCG TTATTAAGAA CAGCAGATT CGAGTGGAA 840  
 AGTTCTAATG CTCAGTTTGT CTTGAGTAC TGCTCTTCCA AGGGGTATAA CATTTCTTGG 900



```

GAACTAGGCA ATGAACCTAA CAGTTTCCTT AAGAAGGCTG ATATTTTCAT CAATGGGTCG 960
CAGTTTAGGAG AAGATTATAT TCAATTGCAT AAACCTCTAA GAAAGTCCAC CTTCAAAAT 1020
GCAAACTCT ATGGTCTCTG TGTGGTCCAG CCTCGAAGAA AGACGGCTAA GATGCTGRAAG 1080
AGCTTCCTGA AGGCTGGTGG AGAAGTGATT GATTCAAGTA CATGGCATCA CTACTATTGG 1140
AATGGACGGA CTGCTACCAAG GGAAGATTTT CTAACCCCTG ATGTATTGGA CATTTTTATT 1200
TCATCTGTGC AAAAAGTTTT CCAGGTGGTT GAGAGCACCA GGCTTGCAA GAAGCTCTGG 1260
TTAGGAGAAA CAACTCTGCG ATATGGAGGC GGAGCGCCCT TGCTATCCGA CACCTTTGCA 1320
GCTGGCTTTA TGTGGCTGGA TAAATGGGCG CTGTCAGCCC GAATGGGAAT AGAAGTGGTG 1380
ATGAGGCAAG TATTCCTTGG AGCAGGAAC TACCAITTAG TGGATGAAA CTTGATCTCT 1440
TTACTGTATT ATTGGCTATC TCTTCTGTTT AAGAAATTGG TGGGCACCAA GGTGTAATG 1500
GCAAGCGTGC AAGGTTCAAA GAGAAGGAAG CTTGCGATAT ACCTTCATTG CACAAACACT 1560
GACAACTCAA GGTATAAAGA AGGAGATTTA ACTCTGTATG CCATAAACCT CCATAACGTC 1620
ACCAAGTACT TGCGGTTACC CTATCCTTTT TCTAACAGCG AAGTGGATAA ATACCTTCTA 1680
AGACCTTTGG GACCTCATGG ATTACTTTCG AAATCTGTCC AACTCAATGG TCTAACTCTA 1740
AAGATGTGGT ATGATCAAAAC CTTGCCACCT TTAATGGAAA AACCTCCGCG GCCAGGAACT 1800
TCACCTGGCT TGCCAGCTTT CTCATATAGT TTTTGTGTGA TAAGAAATGC CAAAGTTGCT 1860
GCTTGCACTC GAAATATAAA TATACTAGTC CTGACACTG 1899

```

## (2) INFORMATION FOR SEQ ID NO:14:

## (i) SEQUENCE CHARACTERISTICS:

```

(A) LENGTH: 592
(B) TYPE: amino acid
(C) STRANDEDNESS: singl
(D) TOPOLOGY: linear

```

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14

```

Met Glu Gly Ala Val Gly Gly Val Arg Arg Arg Asn Gly Ala Glu
 5 10 15
Glu Arg Arg Lys Gly Arg Trp Gly Ser Ala Gly Ser Ala Arg
 20 25 30
Ala Leu Asp Ser Pro Leu Arg Gly Ser Trp Arg Gly Glu Gln Pro
 35 40 45
Gly Glu Pro Lys Met Leu Leu Arg Ser Lys Pro Ala Leu Pro Pro
 50 55 60
Pro Leu Met Leu Leu Leu Gly Pro Leu Gly Pro Leu Ser Pro
 65 70 75
Gly Ala Leu Pro Arg Pro Ala Gln Ala Gln Asp Val Val Asp Leu
 80 85 90
Asp Phe Phe Thr Gln Glu Pro Leu His Leu Val Ser Pro Ser Phe
 95 100 105
Leu Ser Val Thr Ile Asp Ala Asn Leu Ala Thr Asp Pro Arg Phe
 110 115 120
Leu Ile Leu Leu Gly Ser Pro Lys Leu Arg Thr Leu Ala Arg Gly
 125 130 135
Leu Ser Pro Ala Tyr Leu Arg Phe Gly Gly Thr Lys Thr Asp Phe
 140 145 150
Leu Ile Phe Asp Pro Lys Lys Glu Ser Thr Phe Glu Glu Arg Ser
 155 160 165
Tyr Trp Gln Ser Gln Val Asn Gln Asp Ile Cys Lys Tyr Gly Ser
 170 175 180
Ile Pro Pro Asp Val Glu Glu Lys Leu Arg Leu Glu Trp Pro Tyr
 185 190 195
Gln Glu Gln Leu Leu Leu Arg Glu His Tyr Gln Lys Lys Phe Lys
 200 205 210
Asn Ser Thr Tyr Ser Arg Ser Ser Val Asp Val Leu Tyr Thr Phe
 215 220 225
Ala Asn Cys Ser Gly Leu Asp Leu Ile Phe Gly Leu Asn Ala Leu

```

|                                                             |     |     |
|-------------------------------------------------------------|-----|-----|
| 230                                                         | 235 | 240 |
| Leu Arg Thr Ala Asp Leu Gln Trp Asn Ser Ser Asn Ala Gln Leu |     |     |
| 245                                                         | 250 | 255 |
| Leu Leu Asp Tyr Cys Ser Ser Lys Gly Tyr Asn Ile Ser Trp Glu |     |     |
| 260                                                         | 265 | 270 |
| Leu Gly Asn Glu Pro Asn Ser Phe Leu Lys Lys Ala Asp Ile Phe |     |     |
| 275                                                         | 280 | 285 |
| Ile Asn Gly Ser Gln Leu Gly Glu Asp Tyr Ile Gln Leu His Lys |     |     |
| 290                                                         | 295 | 300 |
| Leu Leu Arg Lys Ser Thr Phe Lys Asn Ala Lys Leu Tyr Gly Pro |     |     |
| 305                                                         | 310 | 315 |
| Asp Val Gly Gln Pro Arg Arg Lys Thr Ala Lys Met Leu Lys Ser |     |     |
| 320                                                         | 325 | 330 |
| Phe Leu Lys Ala Gly Gly Glu Val Ile Asp Ser Val Thr Trp His |     |     |
| 335                                                         | 340 | 345 |
| His Tyr Tyr Leu Asn Gly Arg Thr Ala Thr Arg Glu Asp Phe Leu |     |     |
| 350                                                         | 355 | 360 |
| Asn Pro Asp Val Leu Asp Ile Phe Ile Ser Ser Val Gln Lys Val |     |     |
| 365                                                         | 370 | 375 |
| Phe Gln Val Val Glu Ser Thr Arg Pro Gly Lys Lys Val Trp Leu |     |     |
| 380                                                         | 385 | 390 |
| Gly Glu Thr Ser Ser Ala Tyr Gly Gly Gly Ala Pro Leu Leu Ser |     |     |
| 395                                                         | 400 | 405 |
| Asp Thr Phe Ala Ala Gly Phe Met Trp Leu Asp Lys Leu Gly Leu |     |     |
| 410                                                         | 415 | 420 |
| Ser Ala Arg Met Gly Ile Glu Val Val Met Arg Gln Val Phe Phe |     |     |
| 425                                                         | 430 | 435 |
| Gly Ala Gly Asn Tyr His Leu Val Asp Glu Asn Phe Asp Pro Leu |     |     |
| 440                                                         | 445 | 450 |
| Pro Asp Tyr Trp Leu Ser Leu Leu Phe Lys Lys Leu Val Gly Thr |     |     |
| 455                                                         | 460 | 465 |
| Lys Val Leu Met Ala Ser Val Gln Gly Ser Lys Arg Arg Lys Leu |     |     |
| 470                                                         | 475 | 480 |
| Arg Val Tyr Leu His Cys Thr Asn Thr Asp Asn Pro Arg Tyr Lys |     |     |
| 485                                                         | 490 | 495 |
| Glu Gly Asp Leu Thr Leu Tyr Ala Ile Asn Leu His Asn Val Thr |     |     |
| 500                                                         | 505 | 510 |
| Lys Tyr Leu Arg Leu Pro Tyr Pro Phe Ser Asn Lys Gln Val Asp |     |     |
| 515                                                         | 520 | 525 |
| Lys Tyr Leu Leu Arg Pro Leu Gly Pro His Gly Leu Leu Ser Lys |     |     |
| 530                                                         | 535 | 540 |
| Ser Val Gln Leu Asn Gly Leu Thr Leu Lys Met Val Asp Asp Gln |     |     |
| 545                                                         | 550 | 555 |
| Thr Leu Pro Pro Leu Met Glu Lys Pro Leu Arg Pro Gly Ser Ser |     |     |
| 560                                                         | 565 | 570 |
| Leu Gly Leu Pro Ala Phe Ser Tyr Ser Phe Phe Val Ile Arg Asn |     |     |
| 575                                                         | 580 | 585 |
| Ala Lys Val Ala Ala Cys Ile                                 |     |     |
| 590                                                         | 592 |     |

## (2) INFORMATION FOR SEQ ID NO:15:

## (i) SEQUENCE CHARACTERISTICS:

|     |               |              |
|-----|---------------|--------------|
| (A) | LENGTH:       | 1899         |
| (B) | TYPE:         | nucleic acid |
| (C) | STRANDEDNESS: | double       |
| (D) | TOPOLOGY:     | linear       |

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| AAA | CGC | AGC | AAG | GAA | GTA | GGA | GAG | AGC | CGG | GGA | GGC | GGC | GGC | GGG | 3   |
| TTG | GAT | TGG | GAG | CAG | TGG | GAG | GGA | TGC | AGA | AGA | GGG | GTG | GGA | GGG | 48  |
| ATG | GAG | GGC | GCA | GTG | GGA | GGG | GTG | AGG | AGG | CGT | AAC | GGG | GCG | GAG | 93  |
| Met | Glu | Gly | Ala | Val | gly | gly | Val | Arg | Arg | Arg | Asn | Gly | Ala | Glu | 138 |
|     |     |     |     | 5   |     |     |     |     | 10  |     |     |     |     | 15  |     |
| GAA | AGG | AGA | AAA | GGG | CGC | TGG | GGC | TCG | CGC | GGG | GGA | AGT | GCT | AGA | 183 |
| Glu | Arg | Arg | Lys | Gly | Arg | Trp | Gly | Ser | Ala | Gly | Gly | Ser | Ala | Arg |     |
|     |     |     |     | 20  |     |     |     |     | 25  |     |     |     |     | 30  |     |
| GCT | CTC | GAC | TCT | CCG | CTG | CGC | GGC | AGC | TGG | CGG | GGG | GAG | CAG | CCA | 228 |
| Ala | Leu | Asp | Ser | Pro | Leu | Arg | Gly | Ser | Trp | Arg | Gly | Glu | Gln | Pro |     |
|     |     |     |     | 35  |     |     |     |     | 40  |     |     |     |     | 45  |     |
| GGT | GAG | CCC | AAG | ATG | CTG | CGC | TCG | AAG | CGT | CGC | CTG | CCG | CCG | CCG | 273 |
| Gly | Glu | Pro | Lys | Met | Leu | Leu | Arg | Ser | Lys | Pro | Ala | Leu | Pro | Pro |     |
|     |     |     |     | 50  |     |     |     |     | 55  |     |     |     |     | 60  |     |
| CCG | CTG | ATG | CTG | CTG | CTC | CTG | GGG | CGC | CTG | GGT | CCC | CTC | TCC | CCT | 318 |
| Pro | Leu | Met | Leu | Leu | Leu | Leu | Gly | Pro | Leu | Gly | Pro | Leu | Ser | Pro |     |
|     |     |     |     | 65  |     |     |     |     | 70  |     |     |     |     | 75  |     |
| GGC | GCC | CTG | CCC | CGA | CCT | CGC | CAA | GCA | CAG | GAC | GTC | GTG | GAC | CTG | 363 |
| Gly | Ala | Leu | Pro | Arg | Pro | Ala | Gln | Ala | Gln | Asp | Val | Val | Asp | Leu |     |
|     |     |     |     | 80  |     |     |     |     | 85  |     |     |     |     | 90  |     |
| GAC | TTC | TTC | ACC | CAG | GAG | CCG | CTG | CAC | CTG | GTG | AGC | CCC | TCG | TTC | 408 |
| Asp | Phe | Phe | Thr | Gln | Glu | Pro | Leu | His | Leu | Val | Ser | Pro | Ser | Phe |     |
|     |     |     |     | 95  |     |     |     |     | 100 |     |     |     |     | 105 |     |
| CTG | TCC | GTC | ACC | ATT | GAC | GCC | AAC | CTG | GCC | ACG | GAC | CCG | CGG | TTC | 453 |
| Leu | Ser | Val | Thr | Ile | Asp | Ala | Asn | Leu | Ala | Thr | Asp | Pro | Arg | Phe |     |
|     |     |     |     | 110 |     |     |     |     | 115 |     |     |     |     | 120 |     |
| CTC | ATC | CTC | CTG | GGT | TCT | CCA | AAG | CTT | CGT | ACC | TTG | GCC | AGA | GGC | 498 |
| Leu | Ile | Leu | Leu | Gly | Ser | Pro | Lys | Leu | Arg | Thr | Leu | Ala | Arg | Gly |     |
|     |     |     |     | 125 |     |     |     |     | 130 |     |     |     |     | 135 |     |
| TTG | TCT | CCT | CGC | TAC | CTG | AGG | TTT | GGT | GGC | ACC | AAG | ACA | GAC | TTC | 543 |
| Leu | Ser | Pro | Ala | Tyr | Leu | Arg | Phe | Gly | Gly | Thr | Lys | Thr | Asp | Phe |     |
|     |     |     |     | 140 |     |     |     |     | 145 |     |     |     |     | 150 |     |
| CTA | ATT | TTC | GAT | CCC | AAG | AAG | GAA | TCA | ACC | TTT | GAA | GAG | AGA | AGT | 588 |
| Leu | Ile | Phe | Asp | Pro | Lys | Lys | Glu | Ser | Thr | Phe | Glu | Glu | Arg | Ser |     |
|     |     |     |     | 155 |     |     |     |     | 160 |     |     |     |     | 165 |     |
| TAC | TGG | CAA | TCT | CAA | GTC | AAC | CAG | GAT | ATT | TGC | AAA | TAT | GGA | TCC | 633 |
| Tyr | Trp | Gln | Ser | Gln | Val | Asn | Gln | Asp | Ile | Cys | Lys | Tyr | Gly | Ser |     |
|     |     |     |     | 170 |     |     |     |     | 175 |     |     |     |     | 180 |     |
| ATC | CCT | CCT | GAT | GTG | GAG | GAG | AAG | TTA | CGG | TTG | GAA | TGG | CCC | TAC | 678 |
| Ile | Pro | Pro | Asp | Val | Glu | Glu | Lys | Leu | Arg | Leu | Glu | Trp | Pro | Tyr |     |
|     |     |     |     | 185 |     |     |     |     | 190 |     |     |     |     | 195 |     |
| CAG | GAG | CAA | TTG | CTA | CTC | CGA | GAA | CAC | TAC | CAG | AAA | AAG | TTC | AAG | 723 |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |      |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|
| Gln | Glu | Gln | Leu | Leu | Leu | Arg | Glu | His | Tyr | Gln | Lys | Lys | Phe | Lys |      |
|     |     |     |     | 200 |     |     |     |     | 205 |     |     |     |     | 210 |      |
| AAC | AGC | ACC | TAC | TCA | AGA | AGC | TCT | GTA | GAT | GTG | CTA | TAC | ACT | TTT | 768  |
| Asn | Ser | Thr | Tyr | Ser | Arg | Ser | Ser | Val | Asp | Val | Leu | Tyr | Thr | Phe |      |
|     |     |     |     | 215 |     |     |     |     | 220 |     |     |     |     | 225 |      |
| GCA | AAC | TGC | TCA | GGA | CTG | GAC | TTG | ATC | TTT | GGC | CTA | AAT | GCG | TTA | 813  |
| Ala | Asn | Cys | Ser | Gly | Leu | Asp | Leu | Ile | Phe | Gly | Leu | Asn | Ala | Leu |      |
|     |     |     |     | 230 |     |     |     |     | 235 |     |     |     |     | 240 |      |
| TTA | AGA | ACA | GCA | GAT | TTG | CAG | TGG | AAC | AGT | TCT | AAT | GCT | CAG | TTG | 858  |
| Leu | Arg | Thr | Ala | Asp | Leu | Gln | Trp | Asn | Ser | Ser | Asn | Ala | Gln | Leu |      |
|     |     |     |     | 245 |     |     |     |     | 250 |     |     |     |     | 255 |      |
| CTC | CTG | GAC | TAC | TGC | TCT | TCC | AAG | GGG | TAT | AAC | ATT | TCT | TGG | GAA | 903  |
| Leu | Leu | Asp | Tyr | Cys | Ser | Ser | Lys | Gly | Tyr | Asn | Ile | Ser | Trp | Glu |      |
|     |     |     |     | 260 |     |     |     |     | 265 |     |     |     |     | 270 |      |
| CTA | GGC | AAT | GAA | CCT | AAC | AGT | TTC | CTT | AAG | AAG | GCT | GAT | ATT | TTC | 948  |
| Leu | Gly | Asn | Glu | Pro | Asn | Ser | Phe | Leu | Lys | Lys | Ala | Asp | Ile | Phe |      |
|     |     |     |     | 275 |     |     |     |     | 280 |     |     |     |     | 285 |      |
| ATC | AAT | GGG | TCG | CAG | TTA | GGA | GAA | GAT | TAT | ATT | CAA | TTG | CAT | AAA | 993  |
| Ile | Asn | Gly | Ser | Gln | Leu | Gly | Glu | Asp | Tyr | Ile | Gln | Leu | His | Lys |      |
|     |     |     |     | 290 |     |     |     |     | 295 |     |     |     |     | 300 |      |
| CTT | CTA | AGA | AAG | TCC | ACC | TTC | AAA | AAT | GCA | AAA | CTC | TAT | GGT | CCT | 1038 |
| Leu | Leu | Arg | Lys | Ser | Thr | Phe | Lys | Asn | Ala | Lys | Leu | Tyr | Gly | Pro |      |
|     |     |     |     | 305 |     |     |     |     | 310 |     |     |     |     | 315 |      |
| GAT | GTT | GGT | CAG | CCT | CGA | AGA | AAG | ACG | GCT | AAG | ATG | CTG | AAG | AGC | 1083 |
| Asp | Val | Gly | Gln | Pro | Arg | Arg | Lys | Thr | Ala | Lys | Met | Leu | Lys | Ser |      |
|     |     |     |     | 320 |     |     |     |     | 325 |     |     |     |     | 330 |      |
| TTC | CTG | AAG | GCT | GGT | GGA | GAA | GTG | ATT | GAT | TCA | GTT | ACA | TGG | CAT | 1128 |
| Phe | Leu | Lys | Ala | Gly | Gly | Glu | Val | Ile | Asp | Ser | Val | Thr | Trp | His |      |
|     |     |     |     | 335 |     |     |     |     | 340 |     |     |     |     | 345 |      |
| CAC | TAC | TAT | TTG | AAT | GGA | CGG | ACT | GCT | ACC | AGG | GAA | GAT | TTT | CTA | 1173 |
| His | Tyr | Tyr | Leu | Asn | Gly | Arg | Thr | Ala | Thr | Arg | Glu | Asp | Phe | Leu |      |
|     |     |     |     | 350 |     |     |     |     | 355 |     |     |     |     | 360 |      |
| AAC | CCT | GAT | GTA | TTG | GAC | ATT | TTT | ATT | TCA | TCT | GTG | CAA | AAA | GTT | 1218 |
| Asn | Pro | Asp | Val | Leu | Asp | Ile | Phe | Ile | Ser | Ser | Val | Gln | Lys | Val |      |
|     |     |     |     | 365 |     |     |     |     | 370 |     |     |     |     | 375 |      |
| TTC | CAG | GTG | GTT | GAG | AGC | ACC | AGG | CCT | GGC | AAG | AAG | GTC | TGG | TTA | 1263 |
| Phe | Gln | Val | Val | Glu | Ser | Ser | Thr | Arg | Pro | Gly | Lys | Lys | Val | Trp |      |
|     |     |     |     | 380 |     |     |     |     | 385 |     |     |     |     | 390 |      |
| GGA | GAA | ACA | AGC | TCT | GCA | TAT | GGA | GGC | GGA | GCG | CCC | TTG | CTA | TCC | 1308 |
| Gly | Glu | Thr | Ser | Ser | Ala | Tyr | Gly | Gly | Gly | Ala | Pro | Leu | Leu | Ser |      |
|     |     |     |     | 395 |     |     |     |     | 400 |     |     |     |     | 405 |      |
| GAC | ACC | TTT | GCA | GCT | GGC | TTT | ATG | TGG | CTG | GAT | AAA | TTG | GGC | CTG | 1353 |
| Asp | Thr | Phe | Ala | Ala | Gly | Phe | Met | Trp | Leu | Asp | Lys | Leu | Gly | Leu |      |

|                                                             |     |     |      |
|-------------------------------------------------------------|-----|-----|------|
| 410                                                         | 415 | 420 |      |
| TCA GCC CGA ATG GGA ATA gAA GTG GTG ATG AGG CAA GTA TTC TTT |     |     | 1398 |
| Ser Ala Arg Met Gly Ile Glu Val Val Met Arg Gln Val Phe Phe |     |     |      |
| 425                                                         | 430 | 435 |      |
| GGA GCA GGA AAC TAC CAT TTA GTG GAT GAA AAC TTC GAT CCT TTA |     |     | 1443 |
| Gly Ala Gly Asn Tyr His Leu Val Asp Glu Asn Phe Asp Pro Leu |     |     |      |
| 440                                                         | 445 | 450 |      |
| CCT GAT TAT TGG CTA TCT CTT CTG TTC AAG AAA TTG GTG GGC ACC |     |     | 1488 |
| Pro Asp Tyr Trp Leu Ser Leu Leu Phe Lys Lys Leu Val Gly Thr |     |     |      |
| 455                                                         | 460 | 465 |      |
| AAG GTG TTA ATG GCA AGC GTG CAA GGT TCA AAG AGA AGG AAG CTT |     |     | 1533 |
| Lys Val Leu Met Ala Ser Val Gln Gly Ser Lys Arg Arg Lys Leu |     |     |      |
| 470                                                         | 475 | 480 |      |
| CGA GTA TAC CTT CAT TGC ACA AAC ACT GAC AAT CCA AGG TAT AAA |     |     | 1578 |
| Arg Val Tyr Leu His Cys Thr Asn Thr Asp Asn Pro Arg Tyr Lys |     |     |      |
| 485                                                         | 490 | 495 |      |
| GAA GGA GAT TTA ACT CTG TAT GCC ATA AAC CTC CAT AAC GTC ACC |     |     | 1623 |
| Glu Gly Asp Leu Thr Leu Tyr Ala Ile Asn Leu His Asn Val Thr |     |     |      |
| 500                                                         | 505 | 510 |      |
| AAG TAC TTG CGG TTA CCC TAT CCT TTT TCT AAC AAG CAA GTG GAT |     |     | 1668 |
| Lys Tyr Leu Arg Leu Pro Tyr Pro Phe Ser Asn Lys Gln Val Asp |     |     |      |
| 515                                                         | 520 | 525 |      |
| AAA TAC CTT CTA AGA CCT TTG GGA CCT CAT GGA TTA CTT TCC AAA |     |     | 1713 |
| Lys Tyr Leu Leu Arg Pro Leu Gly Pro His Gly Leu Leu Ser Lys |     |     |      |
| 530                                                         | 535 | 540 |      |
| TCT GTC CAA CTC AAT GGT CTA ACT CTA AAG ATG GTG GAT GAT CAA |     |     | 1758 |
| Ser Val Gln Leu Asn Gly Leu Thr Leu Lys Met Val Asp Asp Gln |     |     |      |
| 545                                                         | 550 | 555 |      |
| ACC TTG CCA CCT TTA ATG GAA AAA CCT CTC CGG CCA GGA AGT TCA |     |     | 1803 |
| Thr Leu Pro Pro Leu Met Glu Lys Pro Leu Arg Pro Gly Ser Ser |     |     |      |
| 560                                                         | 565 | 570 |      |
| CTG GGC TTG CCA GCT TTC TCA TAT AGT TTT TTT GTG ATA AGA AAT |     |     | 1848 |
| Leu Gly Leu Pro Ala Phe Ser Tyr Ser Phe Phe Val Ile Arg Asn |     |     |      |
| 575                                                         | 580 | 585 |      |
| GCC AAA GTT GCT GCT TGC ATC TGA AAA TAA AAT ATA CTA GTC CTG |     |     | 1893 |
| Ala Lys Val Ala Ala Cys Ile                                 |     |     |      |
| 590                                                         | 592 |     |      |
| ACA CTG                                                     |     |     | 1899 |

## (2) INFORMATION FOR SEQ ID NO:16:

|     |                           |
|-----|---------------------------|
| (1) | SEQUENCE CHARACTERISTICS: |
| (A) | LENGTH: 594               |
| (B) | TYPE: nucleic acid        |
| (C) | STRANDEDNESS: double      |
| (D) | TOPOLOGY: linear          |

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16

ATTACTATAG GGCACGCGTG GTCGACGGCC CGGCGTGGTA TTGTCTTAAT GAGAAGTTGA 60  
 TAAAGAAATT TGGGTGGTTG ATCTCTTTCC AGCTGCAGTT TAGCGTATGC TGAGCCAGCA 120  
 TTTTTCAGG CAAAAGTAAA ATACCTGAGA AACTGCTGG CCAGAGGACA ATCAGATTTT 180  
 GGCTGGCTCA AGTGACAAGC AAGTCTTTAT AAGCTAGATG GGAGAGGAG GGATGAATAC 240  
 TCATTTGAG GCTTTACTCG AGGTCAGAG GGATACCCGG CGCCATCAGA ATGGGATCTG 300  
 GGAGTCGGA ACCTGGGTT CCCACGAGAG CGCGCAGAAC ACGTGCCTCA GGAAGCCTGG 360  
 TCCGGGATCC CCACGCGTGC TCCCAGGCG CTCCTCCCCG GCGCTCCTC CCCAGGCTCT 420  
 CCGGGCGGTT GGATCCCGGC CATCTCCGCA CCTTCAAGT GGGTGTGGGT GATTTCGTAA 480  
 GTGAACGTGA CCGCCACCG GGGGAAAGCG AGCAAGGAG TAGGAGAGAG CCGGGCAGGC 540  
 GGGCGGGGT TGGATTGGGA CCACTGGGAG GGATGCAGAA GAGGAGTGGG AGGG 594

## (2) INFORMATION FOR SEQ ID NO:17:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17  
 CCCGAGGAG AGCAGCATCA G 21

## (2) INFORMATION FOR SEQ ID NO:18:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18  
 AGGCTTCGAG CGCAGCAGCA T 21

## (2) INFORMATION FOR SEQ ID NO:19:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19  
 GTARTACGAC TCACTATAGG GC 22

## (2) INFORMATION FOR SEQ ID NO:20:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20  
 ACTATAGGCG ACGCGTGGT 19

## (2) INFORMATION FOR SEQ ID NO:21:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21  
 CTTGGGCTCA CCTGGCTGCT C 21

- (2) INFORMATION FOR SEQ ID NO:22:
- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 23
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22  
AGCTCTGTAG ATGTGCTATA CAC 23
- (2) INFORMATION FOR SEQ ID NO:23:
- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 22
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23  
GCATCTTAGC CGTCTTCTT CG 22
- (2) INFORMATION FOR SEQ ID NO:24:
- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 23
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24  
GAGCAGCCAG GTGAGCCAA GAT 23
- (2) INFORMATION FOR SEQ ID NO:25:
- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 23
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25  
TTCGATCCA AGAAGGAATC AAC 23
- (2) INFORMATION FOR SEQ ID NO:26:
- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 23
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26  
AGCTCTGTAG ATGTGCTATA CAC 23
- (2) INFORMATION FOR SEQ ID NO:27:
- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 24
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27  
TCAGATGCAA GCAGCAACTT TGGC 24
- (2) INFORMATION FOR SEQ ID NO:28:
- (i) SEQUENCE CHARACTERISTICS:

```

 (A) LENGTH: 22
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28
GCATCTTAGC CGTCTTCTT CG 22

(2) INFORMATION FOR SEQ ID NO:29:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 24
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29
GTAGTGATGC CATGTAACG AATC 24

(2) INFORMATION FOR SEQ ID NO:30:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 22
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30
AGGCACCCCTA GAGATGTTCC AG 22

(2) INFORMATION FOR SEQ ID NO:31:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 24
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:31
GAAGATTCT GTTCCATGA CGTG 24

(2) INFORMATION FOR SEQ ID NO:32:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 25
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:32
CCACACTGAA TGTAATACTG AAGTG 25

(2) INFORMATION FOR SEQ ID NO:33:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 22
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:33
CGAAGCTCTG GAAGTCGGCA AG 22

(2) INFORMATION FOR SEQ ID NO:34:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 22
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single

```



- (D) TOPOLOGY: linear  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:34  
 GCCAGCTGCA AAGGTGTTGG AC 22
- (2) INFORMATION FOR SEQ ID NO:35:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 23  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:35  
 AACACCTGCC TCATCAGGAC TTC 23
- (2) INFORMATION FOR SEQ ID NO:36:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 22  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:36  
 GCCAGGCTGG CGTCGATGGT GA 22
- (2) INFORMATION FOR SEQ ID NO:37:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 22  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:37  
 GTCGATGGTG ATGGACAGGA AC 22
- (2) INFORMATION FOR SEQ ID NO:38:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 22  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:38  
 GTAATACGAC TCACATAGG GC 22
- (2) INFORMATION FOR SEQ ID NO:39:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 19  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:39  
 ACTATAGGCG ACGCGTGGT 19
- (2) INFORMATION FOR SEQ ID NO:40:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 27  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:40  
 CCATCTCAAT ACGACTCACT ATAGGC 27

(2) INFORMATION FOR SEQ ID NO:41:  
 (i) SEQUENCE CHARACTERISTICS:  
     (A) LENGTH: 23  
     (B) TYPE: nucleic acid  
     (C) STRANDEDNESS: single  
     (D) TOPOLOGY: linear  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:41  
 ACTCACTATA GGGCTCGAGC GGC 23

(2) INFORMATION FOR SEQ ID NO:42:  
 (i) SEQUENCE CHARACTERISTICS:  
     (A) LENGTH: 4848  
     (B) TYPE: nucleic acid  
     (C) STRANDEDNESS: double  
     (D) TOPOLOGY: linear  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:42

GGATCTTGGC TCACTGCAAT CTCGCGCTCC CATGCAATTC TTATGCATCA 50  
 GCCTCTCGAG TAGCTTGAT TAGGGCTGC CGCACCACT CCGTGGCTACA 100  
 CCACTGTGCC CAGGCTGGTC TTGAACCTTT GGGCTCTAGT GATCCACCCG 150  
 CCTTGGCCTC CCAAAGTGTG GGGATTACAG GTGTGAGCCA TCACACCCCG 200  
 CCCCCTGTTT CCATATTAGT AACTCACTGC TAGACCACAA GGATGCACTA 250  
 TTTAGAAACAT TTGCAATGGT CCACCTTTCA AATCACCCAA ACATGTTAAA 300  
 GAAATTTGTA TGACTGGGCA TGGCAGAGTG GCTCATGGCT GARTCTAG 350  
 CATTTTGTGA GGTGAGAGC GCACGATCAC GAGTCAAGA GATTGAGCC 400  
 ATCTCTGACG CATGGTGA ATCCCATCTC TACTAAAAAT ACNAAACAAT 450  
 TAGCCGGGGG TGATGGCAGG CCCCCTGATG CCCAGTACT CGGAGGGCTG 500  
 AGGCAGGAGA ATGCGGTGAA TCCAGGAGGC AGAGCTTGCA GTGAGCCGAG 550  
 ATGGTGCCAC TGCACTCCAG CTTGGGCGAC AGAGCGAGAC TCCGTCTCAA 600  
 AAAAAAAAAA AAGAAAGAAA ATTGGTATGA CTGTTGACTC ACACAGAGG 650  
 TCAGGGCGCT GGGGTGGGGT TGAAGATTAA TGTCATGATA CATCTCTAG 700  
 AGAATCTCTT GTTCTTCAA CTCCACGTCT GCTACCATAT TATTACACT 750  
 TTCTGTTAGT TGGTGTGTTA TGTGTGAATT TTTTTCATA TGTATACAGT 800  
 AATTGTAGGA TATGAACCTG ATCTAGTTG CAATACCTAC TATGAGCTTA 850  
 GCTTTTAAGT TGCTTAAGAA TAGGTAGATC TATGCRAATA ATGATAATTA 900  
 TTATTATTAT TTTAAGSAGG GGTCTCAGTT TGTCACCCAG GCTGAGSTG 950  
 AGTGGTGTGA TTAAGSGTCA CTGCAACCTC CACCTCCACG GCTCAATAAT 1000  
 ACCTTCGACG TCAGCCTCCC CAGTAGCTGG AACCAACAGG ACGGGCCACC 1050  
 ACGCCGTGGT AATTTTTTGT ATTTTTTGTG GAGATGGGGT TTCATCATGT 1100  
 TGCCCAAGGT GTCTTGAAT TCCTCGGCTC AAGCAATCTC CCACCTTTGG 1150  
 CCTCCCAAAA TCGTGGCATC ACAGGACATGA TGGCATCACT GGCATCACAT 1200  
 ACCATGCGTG GCGCTGATTT TGCAAAATTAG ATATGACATT CAATATATC 1250  
 TATTTTTATT TGTTCGCTTA TTGCTGGTAC AATCTCAAGT GGAAATACTC 1300  
 AAGGTTTTTG GTGTATTATT CTTACTCAAC CAATATTATT TAGACTCTTA 1350  
 CTAAGCACCA ACATGATCAC ATGCTGAGC TATGGCTAGC ATAGCGTGTG 1400  
 AGACAAACTT AATCTCTGTT TTGCTGGAGC ATATACTTA GTAGATGAAG 1450  
 CCAATTTTGA GCACATCAC ATACTACACA AATTGAGGAT GCTACAGAGT 1500  
 TGTCTAACAA ATTGAGGHTG ACCTGAGAGT GTCTACAAA TTGAGGATG 1550  
 TATGAGAGTG TGTCTAGGAG ACCTGCTGG AGATTAGAGG AAGCTTCTCT 1600  
 TGGAGGAAGT TACATTTTCA CTGAACAACA CTGCCATCTG CTCGAGGTTT 1650  
 TGTAACTGCA TTCACATCCC GATTCTGACA CTTCCACATCC CGATTCTGAC 1700  
 ACTTCACCCA GTTACTGTCT CAGAGCTTGG GTCCGATGT GTAAACACAG 1750  
 GACAGTATGC ACTTTGGCAG GTTGTGAGAA GGGAGAGAAA CACAGTAAA 1800  
 GCACCTGTAT CAGGCATACG ATAGCCTCTA ASGCTGCGAT GCTTGTATG 1850  
 ATTATACATC AGCTTAGACA TCAGAGAAAA GCTGAGAAAA AGCTCGACCA 1900  
 ACAGCGAAGG ATAAATGCCG AGAGGAGAAA TTTGGCAAAG GCTCCAAATT 1950  
 CAGGGGCACT CGTACTCTTA CACTTTGTAT GGGGGCTTCA GGTCTTGAT 2000  
 TCCAGACATT GGAGCACTTA ACCCTTTAAG ATTGCTAAAT ATTGCTTTAA 2050  
 TGAGAAGTTG ATAAAGAATT TTGGGTGGTT GATCTCTTTC CAGCTGCAAT 2100  
 TTAGCGTATG CTGAGGAGAG ATTTTTTCAA GCNAAAGTAA AATACCTGAG 2150  
 AACTCTGCTG GCACAGGAC AATCAGATTT TGGCTGGCTC AAGTGACAG 2200  
 CAAGTGTGTTA TAAGCTAGAT GGGAGAGGAA GGGATGAATA CTCCTATGGA 2250  
 GGTTTTACTC GAGGCTGAGA GTGATACCCG GCGGCATCAG AATGGGATCT 2300  
 GGGAGTCGGA AAGCTGCGGT TCCACACAGA GCGCGCAGAA CAGCTGCGTC 2350  
 AGGAAAGCTG GTCCGGGATG CCGAGCGCTG CTCCCGGGG GCTCCCGCT 2400  
 GGGCGCTCTT CCGAGCGCTT CCGGGGCGCT TGGATCCCGG CCACTTCCCG 2450  
 ACCCTTCAAG TGGGTGTGGG TGATTTCTGA AGTGAAGCTG ACCCGCACCG 2500  
 AGGGGAAAGC GAGCAAGGAA GTAGGAGAGA GCCGGCGCAG CGGGCGGGG 2550  
 TTGGATTGGG AGCAGTGGGA GGGATGCAGA AGAGGAGTGG GAGGGATGGA 2600  
 GGGCCAGATG GGAGGGGTGA GAGGCGTAA CCGGCGCGAG GAGAGGAGA 2650  
 AAGGGCGTAG GGGCTCGGCG GAGGAAATGG CTGAGACTCT GAGCTCTCG 2700  
 CTGGCGGACA GCTGGCGGGG GAGCAGGCA GGTGAGCCCA AGATGCTGCT 2750

GCGCTCGAAG CCGTGGCTGC GCGCGCGCGT GATGCTGCTG CTCCTGGGGC 2800  
 CGCTGGGTCC CCGTGGCTGC GCGCGCGCGT CCGGACCTGC GCAAGCAGAG 2850  
 GAGTCTGTGG ACCTGAGCTT CTTACACGAG GAGCGCTGC ACCTGCTGAG 2900  
 CCGCTGCTGC CTGTGCGCTCA CCAATTGACGC CAACCTGGCC ACGGACCCGC 2950  
 GGTTCCTCAT CCGTCTGGGG TAAGCGCCAG CCGCTGGGTC CTGTCGCCCT 3000  
 TCGTCTGCTC CTGACACCTA TGCTGCGCCC GCGCAGGGCT CTCCTCTTTT 3050  
 TGGCGGGGAAA CAACCTTCACA CGGGACCTGC CCGCGCTGTC TCTCCCCACC 3100  
 CCACTTCGCG CCGTCTGCTC TCGCTCTGCC TCCTCTACTC TCAGACGCTA 3150  
 AACCGCTTTT TGGCGGGTAT CATTATAAAA ATGATTTTGA GGGTTACGAT 3200  
 TGCAGTCTGT TTTCCAGGGT ATATTGCAAT GTGGTGGCAT CTGGGCTCTT 3250  
 AGTGTAACTG TCACCCGAAT GTTGATCATT GTATCTAATA GGTAAATTCT 3300  
 CATCCCTCAT CCGCTCTCCA CCGTCCCACT TTTTGGAGTC TCGAGTGTCT 3350  
 ACATATTCCAC TAAGTCCAGT TGTAACACATT GTTTAGCGCC CACTCTAAT 3400  
 GAGCTTTTTT GTCTGATCTA TCTGTAGAT GTTGAATAGG CACACCTAAT 3450  
 GGTGAGGTAT AAGTGGAAAT TCGAAAAGAA AACTGCCBAC TTGCCCAAGT 3500  
 ACTTCCCTAG CCAAGAGGAG GAAACCCAGG CAGGTGCACC TGAAGGCGCT 3550  
 TGAGTGTCTG ATTTGCTGTG CAGTGATAGGA CAAGTAAGAT TGTGCATAGC 3600  
 CTCTGTATTT TAAGACTGTG TTAGGAAGAT TCTCTTTCT TTTCTTTTCT 3650  
 TTTTCTTTT TTTTCTTTT TTTTTTTTTA GGCAGATGAA AAGGGGCTCA 3700  
 CAGACAGGA ATAAATATCT AATATATCAA TAATGTAGAC CTAGGAGACT 3750  
 ACTGCAGTGA CTTACAAGT CTTAATAAAA AGATGTCTCT CCAAAATGGG 3800  
 GCTGCAAAAT GTGGTGTCTG CTTATCAGCT CTAAGTTTTT TCTTACCTG 3850  
 AGAAAGAAAG AACCTGATGC AGGTTCAAGG CTCTCGCCCC ATGAATGCAG 3900  
 GCTGACTCCA AGATGGGGAG CTACAGGGAC AATCCCAAGT CTTTAGGCC 3950  
 TCTTATTAG GCGCTGGGAG CCGTCCAGGA TGCGCACATC TTGACGAGCC 4000  
 CAGATAGAG GAGACCCGAC CATTATCTCA CCGTGTGCTG AATACCTAG 4050  
 ATGCTGTCTT CCGTTCAGCC ACACTATAGT TGCAGCGCT AATTAAATGG 4100  
 GTAGTGTACT GGTAAAGAA TGACAGAGCC ATCTCGGCTT GACTCTCAGC 4150  
 TCTGGCAAG ATGAGTGACT TGGTTTTTCC ATATCTCTTG GCCACACCAA 4200  
 CCGTGATTTC TTAGCTGTGA GAATGGAATT TCTCAAGCTT GCGTCAAGGA 4250  
 TATGTCGCG AGGATGAT GATATGGTAA GAGCTTCTCA GTGTGTAGC 4300  
 CATAGTAAT CATGACGCT TCAAGAGAT TGTTCCTTG TAGAGACTAG 4350  
 TGAGCATTTG GTAGCCATTC ACCGGTTTTT TGTTCCTTG GATCATAGTT 4400  
 AACCTCTCT TTTCTCTCTG GCACTACAAT TTTCTGGTGG GGAAGAATCC 4450  
 TTTCTCTCTG CCGTCTCCCT TAAGGATAGG AAGCTGATAC TAGGACAGAA 4500  
 CTAGTGGGG GATGAGGAGA TTGTTCCAGA GAATGCTGA ACCATAGGGC 4550  
 TCCAGATCAC AGGACCCGAC TTTAGCTGCT TTTAGCTGCT GGGTGGGGG 4600  
 GCGCGGTTAC TGACATGGG TATGAGTAG ATGCTCATTT ACTGAAATGT 4650  
 GAGGACCTGA GGCCTCTCT ATTGCTGTAG CCAGCATATT CCCCCACCTC 4700  
 TCCCCAAGAA AGGACAGATG GGGGTTCCCC CCGTGAGTAA CAGGTCCAAA 4750  
 AGAAAAAACA TACAGTGGGA CTTCCAGGAT CTGGGCGCTGA TCACCCAGCA 4800  
 GTACAGCTCC CGGCAATTGA CTAACACCCC CTTACACAGT AGAATTTCCA 4850  
 ATCTGCAATT TATGAGAGT GAGACCTTTA TTTCTTTTAA ATACATCTCT 4900  
 CTATTTCACA GAGCACCTTT TTTTCCCCCT CTTGCACTT TTTTGTAA 4950  
 GACTGGAGTA TAATGAATA CCAAGAGAGC ATACATGTGT ATACATAAAA 5000  
 CTTTTTTTCT GGTTTACAAA ACAGTTTATT CTTGTCATA CGTGTCTCTC 5050  
 TCAAGGCTGT GCTGCTGTCT GTTCCAGCCC GCTTCGCTGT GAGAGGCCAT 5100  
 CTGCCATCC TGCTGCCGAC AGGATATGAC AAGCACACC AGNSTGTAT 5150  
 CTGCTAAGAC CTAAGAGAGG GAGGAACCCC CTTCTCAT CTAGACACTA 5200  
 GCTCTTAAT TAGAGTGTGA GGGTCCATCT CCCCAGGAG GGCACAGGCC 5250  
 CCAACAGGCC CAGCCATCTC AGAAGACAA CTAAGCTTT GTAGGGGCTC 5300  
 AAGTATAGAG AGAGTAAAG GCGTGTGTT TAATTATTA CAGTTCCTCA 5350  
 CAAGTAGAGA TGTTGGGGCG GGGTGGCAAG GCTGAGCAG ACGAAGAGT 5400  
 AAGGAATAG GAGGAGAG AGGACAAAA CAGGTGAC CTTCTCACT 5450  
 CTGTGCAATT CCGTGGCCCT GTTCTAAGCA CTTCTAGT ATTAACTCAT 5500  
 TTAGTCTTGG CTACAACACT GTGAGTAAGT AGTTTGTGCA CCCCATTCT 5550  
 AAAAATGAAG ARAGTGAGGC TCAGGGAGGT TAAGTAACCT GGGCACAGTT 5600  
 TGAACATAGA CTTGATCAT ATGAGATAAT AGTGCCCAT AAAAAAGAAA 5650  
 GCGATATTA TTTTAAAGAG GARAAGAGAT AGGATATGCT AGAAGAAGAT 5700  
 TTTTGTGAAR GGAATGAGT GATGATATA ATGAAAAGAA CAGTACAT 5750  
 GAGAGTAACA GTATCAGGCC CCAACCTCTC ATCTAAGTA CTTCAAAGAG 5800  
 GCGTAAGCAA ACTTAGCTAC TGGCGTGGTT CTAGTCTCCA TGATGGCAAA 5850  
 TACATTGTGT ACAGGCCAAC TCCACAAAA ACTTAATAC CAATGATAGA 5900  
 CGAATCTAAA ATTTAAAGAA AAAAATCTTT CATTTGTGCT TCTTCCAGAA 5950  
 GGGACTATTA CAGAAAGAG ATCAAAATC TTTCTAGGCT TTTCTGCTG 6000  
 CAGACTTCCA AAGAGAGCCC AGCCCTAAT ACACACTGTC CAACTGGAA 6050  
 ATAAATATA GTGGGCTCTA TATGCAAGGT CATATGTAAT TTTAAATTTT 6100  
 CTAGTAGCCA TATTAAAAAG GTAAAAAGAA ACAAGTGAAA TTAATTTTAA 6150  
 TAATTTTATT TAGTCAATA GATCCAAAT GTTTCCTCAG CATGTAAATCA 6200  
 ATATAAAT ATTAAGGAGT TGTGTATTT TAGTAGAGAC AAGGTTTAC 6250  
 TTTCTATA TTTCTGAT ATATTATTTA CAGCTTCTCA GATATATTT 6300  
 CTTCTCTCT TTTTCTTCT CGAGACAATT TGCTCTTGT CACCAAGCT 6350  
 AGAGTACAA TGGGTTACCT CGGCTCACTG CAACCTCGCG CTCGCCGGTT 6400  
 CAGCTATTT CTCTGCCCTA GTCTCCCAAG TAGCTGGGAC TAGAGGATCA 6450  
 CACCACAGC CCGTGGTAA TGTGTATTT TAGTAGAGAC AAGGTTTAC 6500  
 CATGTTGGC CAACCTGCT AGCTAGGCT ATAGGCCAC 6550  
 CTCGGCTCC CAAAGTGTG GAAATTACAG CGTGAAGCC TGCACCCGCC 6600  
 CTCAGATTA CTATATTTCA AGCGTTCAGT AGCCACATGT AGCTAGTGCT 6650  
 ATGAGTAGTG ACAGTACAGA TCTGCAATTC AATTAAAGCA CGTATACAG 6700  
 CATAGTTCAC TAATGCAGCG TAAAAAAG TATAGTCTG AGTCGTTGTT 6750

0098813.11901

AGAAATCTCTA AATACCTGCAG AGCAAAAGTG GTACGACACG CAATCTCAGT 6800  
 GATAATGCAA CCATGCTTGC TTTTCATTGC AATTGCTTGA TTTTCCCTCA 6850  
 CGAAGATTCA TCATTTTTTG CCAATTCRAAT AAATATTTCAT TGTAAAAACG 6900  
 TTTCAATAT AGACTCTGC ATCTTCATAG ACAGAGTCTG TTTTCACATG 6950  
 TAGAAATAAT CTTATCAATG TTAACACACAC GTTTTGATAA CCAGGTGTGG 7000  
 AAAGAGGTGC AGACTCCCCA GTGCGCTATT GATGGCAGAA ATATTCCACG 7050  
 CCAAGAGGGA ACARAGGGCT GGGGACAATC ACACACTCTA TGTCTCCTAA 7100  
 CTCCTGGGAA GTGCTGTCCC TCTGATTGAG CTCTATATAT TGCCCTCCOC 7150  
 ACTAACCTTG TCGACTGTGC GTTGCGAGCC TTTGCAAGGT TACTGCTCTT 7200  
 GTCCTCTCA CAGAAATATC CTTCTACTCT CTCTGCTGAT TCAACTCTG 7250  
 GCTATTCTCT GATGACACTG TCTTCCCTGT AGCCCTTTTG AGTAATGGCT 7300  
 GCATATTCTC CCATAGTCCA GTTCTTTTCC TGTCTCCAG TCTGGCTTCT 7350  
 GGTAGCACAG CCACTAGTIT GAACCTCCATA CTCTATAGT TCAAGTCCCT 7400  
 TTTGACTTGT TACCTTGGGC AAMATTACTC CTTTTGTCTA GGTTCCTTGT 7450  
 TTCTAAATG ACAGATANTA TCCATCTGCG TCCATGCTGT TTTTGTGAA 7500  
 TTGAGTGAAA GAAGCGGGCT AGCTTCCCTA CAGCTCTAGT GTAGACTAGC 7550  
 CTGATGTGCA TTACGGGTGA TGCCATGACT CAGTGTGTT TCTCATCTC 7600  
 CACATCTGGC TCTCATCCAG TGCTCCTGTG TACGGCACTC TGTCCCCCTC 7650  
 TTACTTACTC CCCTTATTA ACTGAAGACT GGCATGATC TCACAGTTTC 7700  
 CTCCTCACTT CTTAGTCTCA CCATCATCTT AGATGACTCT AAGTCACTTA 7750  
 GATAAAGTGT CTGAGTTCTT TCACTCAAT TTTTATTATA CAGATATGT 7800  
 TACACTCAAG TTGTACRGA ACCAGCTTAT CCAGCTCATG AAATGTATGC 7850  
 ATTTTCATCT AACTCTGTAT TCACTGACAT CCTGTGGTA TCTGGAATC 7900  
 AGCCATGGTG AGAATATTTA CCATGGAAAT TGGCAATAC TAAAAAGCAG 7950  
 AGCACCTTTT TTTCTGAGAG CAGACACATA GCTCTCTAC TCCATAGCAC 8000  
 CCACTATACG AATTTTAAAT TCCCTTCACT GACAGCTCT TCTCCTCTC 8050  
 TACTTCTTC ATATCTGATT TGAAGTCTCT AATTATCTG TCTGACCTCT 8100  
 CTTGTAATAA TAACCCCAAA TCCCTGTCCC ATTTGTTCTC CTGCTAAAT 8150  
 ACTAAACCTG GTTAGTCCA ACCATATTTT CTCTCTTGG AATCTCAGG 8200  
 GTGGCCCAAA AACCTGGAAA TGGAAAAATA TTACTTATTA ATTTTAACT 8250  
 RTATTAATAA GCATTTTAA TGTCTTATT CCAGTCTCAG TGGCCACCT 8300  
 GTATGCTGG GCTATGAGC TCTTGCGGA GAGAGGAGTG GACATCTCT 8350  
 CAGCAACRGA GACTGATGTT GCACCAACCA TTTTGTAGT TCCAGACTT 8400  
 CCTGGCCCTT AGTGTACCC TTAACTCTCC ATTCTCTGC CTTCACATT 8450  
 CTCTACTTTT TAAATATCTC TGACTCCACC TTAACCTTAT CATCTTACG 8500  
 ACATGACCAT ACTTCTGCTT CCCAAGAAA RTAGACRAAT ACTTCTCTT 8550  
 CTTTTTCTCT CTGTCAITCA ATCTGACAG ATGTGATGCC TAAGTCCAG 8600  
 TTTCTCTCT CTCTGATCT CAGCTCTCT CACTCTCTA TGGCCTGAT 8650  
 CCCCTCCCTT CCCCACGCC CAAGGACTTC GCTCTATCAG TCACCTTCT 8700  
 CCTCTCCTGT ATCTTCAACT CCTCCCAATT TACTGGCTC TCCCTCAAG 8750  
 CTTTCCCAAA GCCTTTCCCA TCTCAATTAC CTCTCCGAC ATGCCCTTGC 8800  
 RGAACACACC CCCTTTCTCT CTTCCCTCTG CGAGCTGTGT CTCTCTGTCT 8850  
 TGCCCTCATG ATGCAACAT CATCTGTCA CTAAATACA TCTCTCCAG 8900  
 ATCATCAATG GCCTTCTCTT GTTGGGAAC CTAAATAACA CTTTATCTTA 8950  
 TTTGGTCTT GTTATGGGTT GAATGAGGTT ACCCGAAAT CCATATTAGA 9000  
 AGCTCTAACC CCAGTACCT CAGAATGTGA CTTTATTGG GAATAGGCTC 9050  
 ATTGCAGACG TTATTAGTTA GGATGAGGTC ATACTGGAAT GTGATGGGCT 9100  
 GCTTATCTAA TATAGCATGT GTCTTATATA CAGGAGAAA TTTGGAGACA 9150  
 GACACACRGA TAGGAGATAT ACCATGTGAT GACAGAGTT ATGAGGTTG 9200  
 AGTCAAAAGC CTATGGGAAC TTAGGAGAAA GACCTGGAAC AAATCTCTTC 9250  
 CTGGCGCTAG AGAGGGAGTA TGGCCCTGCC ACTACTTGA ATTCACAGTT 9300  
 TCGGCTTTTC AAACCTGTAA GACATATAG TTTCTGTGTT CAACCAATT 9350  
 GTTTTGGACT ACTCTGGCAG TGGCGCCCTA RCAACRAAT AGAGTCTTT 9400  
 GAGGCAATT GCAAGGAGG ACARAGGAGT CACTTCTTA CAGTCTTAG 9450  
 TCTCTCGCTT TCTTGTGTG GGGGTGTTT CTAAACATCT CTCTCATCT 9500  
 CTCTCTCTCT AGTTTGTCTT AAACATTGGT GTTCTCAGA CTTCTGACCT 9550  
 AGGCCTTCTT TCACTTCCAC ATATTCCCTT GGGTGGTCTC ACCCACTTCC 9600  
 AGAAATTAAT TAAATTACTG CTCTAGCAGT ACTGTGCTGG AACTGTGTTA 9650  
 RCAACTGGCT CTCTGGGAG AGGGGAGACT GGTGATGCT TTTTGTGAT 9700  
 TTCTGTGGTG TAAATATCTC CTCTCGAGGC ACTTCCACAT TCCACAGT 9750  
 TTAACAGCTC GCTCACARAT TTTCTCCAAA TTTAACATT GGTCTTCA 9800  
 GGCCACACAC GTGGTACAGC CAACCTCCAG ACACCTCTGC TTTTGTGTA 9850  
 GAGAGAGTA ACTATTITTT GTACAAAAGG TAAATATAA ACACCTGCAG 9900  
 GCCCCTTTT TTTCTTAACT AAATGTCTCT AGAATNGAA TAGCTGAGC 9950  
 TCTTTATAG CTATTAATCTG TTATTTCAT GTACATGAG TGGTGGATT 10000  
 ATTTTCTCTT TATTTTCTT GTATATGGTT GAATACACT ACCTTTGATC 10050  
 AGTTTATGTT TTATGGCATG TTTTGCACCC ATATTAACT TAGTTTGTG 10100  
 CAGAGGGCCT CAAATATTAT TTTCTCAAAAC AAGAAAATAT TTAATTGCAA 10150  
 AGGAGACAAA CAAAAGGGA CTATATACCA AAATCTTGA ATGTGATTTC 10200  
 TTGTACTGCT CAGTGTCCAA GTGGTAAACC CARACAGAT TGGGTTTCA 10250  
 TTTTGTACG GAGCTCTTT GTCTGGAGC GACTTAACCT TACATCAGC 10300  
 GGGCCTTGCT CATTCATCA CTTAAGTATT TATTAAACAC CAGCGGTGTG 10350  
 CCAAGTACTT ATCTAGTAT CGGATAGATT CTGATAGTC AGTCAGGTCC 10400  
 CTGCTCTCAG GAGGCTTGCA GCAGAGATGG GGGCTGCAT AGAGAGTAG 10450  
 CCAAGGAAT GAAAAGGGA GTTGATTTC GAGAGTAGT AATGCTATGA 10500  
 AGAATATGA GAGCACTCT ACTGAGGAG ACTGACCCA GTGGTACG 10550  
 TTTGTACCT TAAGACACAG ACTGTGACC AGGTGACTCA CAGATGCCG 10600  
 TCAATGTATG CCACAGCAC TTTTCCAGGT GCTCGTTTTC TCCACTTCTC 10650  
 CAGTCTCTG CCGACGGCG ACTGCTTACA AATACAGTA GAGGAATCTA 10700  
 AATGAGGTC TCTCATATC AAACCCAAAT AATAAGCAA GGAACAGAT 10750

|             |             |             |             |             |       |
|-------------|-------------|-------------|-------------|-------------|-------|
| ATCCGCTCGT  | CTGAGGCAC   | TGGAAACAGC  | CCAGCTCGGA  | GTGTGTCTCT  | 10800 |
| CTGAGGAAGT  | TCTTCATCTT  | TTGTTTAGGG  | CCATACGTAG  | GACCTGTGTA  | 10850 |
| CTGAGGGGTT  | CGCACTCGCT  | GACATTTCTA  | CTGAGAGCTC  | GCGTGTCTAT  | 10900 |
| ATTCGCGCGC  | TGATATGTCT  | TGCTGAGTCT  | CAGACACAGA  | GGCGGAGGCG  | 10950 |
| TACCAAAATG  | GGAAGATATG  | TTGTTGTGCT  | TGATCTGGTA  | TCTCAAAATG  | 11000 |
| ATCAACCTGT  | CTCTCTACGC  | CAAAATGATG  | TACCATCTAT  | CAGAGGAGCTT | 11050 |
| CCGAGAGTAT  | ATCCTCTGACT | CTCTCTCAAC  | ATCCAATATA  | TAAATCAATC  | 11100 |
| TAGGCAGAGT  | ACAATATGCT  | AGGCCATATA  | TCCGACGACT  | TTGGGAGGCT  | 11200 |
| CTGGCAGGTG  | GATCATTTGA  | GGCGAGAGCT  | CTAGACAGCG  | CTCGGCCAAC  | 11250 |
| ATGTTCTGCT  | CTCTTAAGAT  | TAAAAAGAGT  | TTTCTTAAAA  | AGCTCAAACT  | 11300 |
| ATGCTGCTGT  | CTCTCTAGCT  | AGCTCTGAGT  | TTCTTGCTCT  | GCTTACAGCT  | 11350 |
| CTGAGAGTCT  | TTGTTCTTAGT | CTGTGGAGTG  | CAGCTGAGCT  | CTCTFACAGT  | 11400 |
| CTTCATGTGT  | TCAACGGTGT  | CTCTATATTA  | AGTTTGGACT  | CTGGAGATCT  | 11500 |
| CTCTCTAGCG  | GGCCTCTCGT  | TAGGTGGCGC  | TATTTGCTCT  | TCGACTTCGT  | 11550 |
| TCTCTGATGT  | CTGCTGAGCT  | TCTCTCTAT   | CTCTCTGCT   | TGGAGAGACT  | 11600 |
| GTCTGCTCAAT | GTGGCGAGCT  | TTTCTTCAAT  | CTGATCCGAC  | TGCGTTTAAAT | 11700 |
| TTTCCCTCTA  | CTTGAATAGG  | TTTGAATGCT  | ACGAGTGAAC  | CCATTTCCCT  | 11750 |
| TGAAAATAGT  | CTTCTCAACA  | TTCCGACTAC  | GGCTTAGAGT  | TTTCTAGATT  | 11800 |
| AGGAGTATCT  | CTGAATGTCT  | CCGATGAGATT | TTTAAAGATAT | TGCGTCAATTA | 11850 |
| CTGCTCTGCT  | TTTCTTCTGCT | TTTCTTCTGCT | ATCTAGACCA  | TCTGCTGGAG  | 11900 |
| AGCTAGCCTGT | CACAAATATAT | TGTTGTAAATA | TAATTAACAGA | TGGCTTTATC  | 12000 |
| TCTTTAAGTA  | TAATCTTCTCT | TTTCTCACTA  | TAAACACAGA  | GCACAGCGAG  | 12050 |
| TTGCTGTGGT  | GCCCATCTGCT | GTAATCCGAG  | CAGCTTTGCA  | GGCTGAGGCT  | 12100 |
| AGAACTGCTG  | CTGAGTGCAG  | CTGCTTARGA  | CAGCTCTGGC  | CAGACTGGGT  | 12150 |
| AAACCTGCTG  | CTGCTGAGCT  | AGGAGGAGCT  | AGGCAACAGA  | ATCCTGCTCA  | 12200 |
| TCCGTATAGT  | CCGCACTACG  | AGGGGCGGAG  | ATCATGCGAC  | TGTACTTCAG  | 12250 |
| CCGAGGAGAG  | AGAGGTGGCA  | TCTCAAAACA  | CACACATACA  | CACACACACA  | 12350 |
| CTGGATGATG  | AGAGACCCCA  | CACACACACC  | TAATTTGATA  | TAATTAATAAT | 12400 |
| CACACACACA  | CACACACACA  | ATATAATTAAC | TATATCAATA  | TATATCAATA  | 12450 |
| CTGTACCATC  | TAGCTACGAC  | TATATTAATAT | TATATGATTA  | TATATCAATA  | 12550 |
| ATTCTCTCGT  | GATATATATAT | ATATATTTCT  | TTATTTGGTA  | TAATTCAGTA  | 12600 |
| TGAGTACTAT  | TTATTTTCTT  | TACTGTGGCA  | AAATCGGACA  | AACATAATAT  | 12650 |
| TTCTGCCATT  | TAGAGTGTAG  | AGTTTGTGCT  | TAATTCAGCT  | CTACATATGT  | 12700 |
| CTGAGCTCTA  | CTACACATAT  | CTATCTGCTC  | AGCTTTTCGT  | TTCTCCCAAC  | 12750 |
| TGAGCTCTCT  | CTGAGCTCTA  | CTGAGCTCTA  | CTGAGCTCTA  | CTGAGCTCTA  | 12800 |
| TACATTTCTT  | TTTATTTTGT  | GTTTGTACCA  | ACTGAAATAT  | AGCTCTCTCT  | 12850 |
| CTCTTACTTA  | GTTCAGATTA  | GCATTTCCAT  | TTTATTAGCG  | GGTGTGTGTA  | 12850 |
| GTGGCGCATC  | CAGATGCGCA  | TCCCTCTTAG  | AGCTCTCTCT  | GCGGTGTGAG  | 12900 |
| TATTTCACTG  | AGGGTGAATT  | CGGGTTGATA  | ACATTTTAAA  | ATGCACTCTT  | 12950 |
| CTGAGCTCTA  | CTGAGCTCTA  | CTGAGCTCTA  | GTATTTTAGT  | ATGAGTCCCA  | 13000 |
| CTTACAACAA  | TAGGATGAGAT | TAGGATGAGAT | TAGTTTTATC  | CTTAGACATC  | 13100 |
| AGAGTACTAC  | TTTGTCTCAAT | TTTCTATCTT  | TATACATATT  | TAATTAAGCA  | 13150 |
| AGAAATATAT  | TGAAAATTTT  | TACATAGCAT  | TCAGATCTTT  | AATTCCTAGG  | 13200 |
| TGAGCATCAT  | TTAAAATCTT  | AAATTTCCGA  | TTGTTGGTGT  | AGCCGCTATA  | 13250 |
| TTATTTAGAG  | TTGGGAGGAT  | TTGTTGGGAG  | TTGTTGGTGT  | AGCCGCTATA  | 13300 |
| CTGGAGGCTG  | CAATGCGCTG  | AGATCAGCCC  | ATGCTCATCT  | AGCGTGGATG  | 13350 |
| ATGAGATAGT  | AATCTGTGCT  | CAAAAAAANA  | AAAAAARAAA  | AAAAAGAGAA  | 13400 |
| GAGGAAGATAT | TGGCAATCAT  | TGCTCTCAGCA | ATTAATTTCT  | GACTGTGAAT  | 13450 |
| ATGACTTCAT  | TGACCAAAAC  | TAAATAGGCG  | ATTTCCAGAG  | TGCTFAGCAT  | 13500 |
| ATGCTGCTCT  | TTAGTGGAGT  | CTGATGAGCT  | GTGCTGAGCT  | GGCAGCATTG  | 13550 |
| CATGTTTGTAT | TTAGTTCAGAA | CTGATGAGCT  | GTGCTGAGCT  | GGCAGCATTG  | 13600 |
| TTTCTGTTTCT | TTAGTTTTCG  | TGATCTAAAT  | ATTTATGRAC  | ATTCCTCATC  | 13650 |
| TTTGTGAGGA  | AGGGATATATA | GATCATTTCT  | ATTTCCATTT  | CTAGACATCT  | 13700 |
| GTATCACTCT  | AGACATGACA  | TAGGCACCTA  | TTTGAGCATC  | TTTGTGCTGT  | 13750 |
| CAGAGATATC  | ATCATGAAAT  | TTGTTCAATT  | AGAGGTGTCA  | GTGATGGGAA  | 13800 |
| TTGCTGCTCT  | TGAGGAGGAT  | TGAGGAGGAT  | AGAGGATATC  | GAGAGTATGA  | 13850 |
| TGAGAGCTCT  | AGTCAATTTA  | AAAGTCTCTC  | GGAAGAGCAT  | GTGTGTGAAG  | 13950 |
| CTCATTTTAA  | AGTCAATTTA  | CCAAATTTCT  | TGTTGTGTGT  | TTTGTGTGTT  | 14000 |
| TTTAAAATAG  | TTATTTTCTT  | ATAGAGCAGT  | TAGGTTTACA  | GCAAAATTTA  | 14050 |
| TTTCTGCTCT  | TTATCATATC  | ATAGAGCAGT  | CTCTCCAGCA  | ACATGATGAT  | 14100 |
| GTGAGTCTCT  | TGATCATATC  | ATAGAGCAGT  | AAAGTCCATG  | TTTCCAGGAC  | 14200 |
| AGGACCTCAT  | TGACATACAT  | TTATCAGCTT  | TGAGCAACAT  | TATATAT     | 14250 |
| TTGCTGCTCT  | CTGCTGACAT  | TCTATGGGTT  | ATTTCTCAGT  | CCCTGGAAT   | 14300 |
| TTGCTGCTCT  | ATTTATATCA  | ATTCCTGCTC  | CTCTGCAATT  | CCACCCCGAG  | 14350 |
| TTTCTGCTCT  | AGACACAGAT  | CTGCTTCTCT  | CTCTTCTTCT  | GGAGTATTTA  | 14400 |
| TTTCTGCTCT  | CATATCAGCA  | TAAAGGCTTA  | TGAGTTTCTT  | TAAAGTTCTT  | 14450 |
| CACACACCTC  | TACTTGTCTA  | TATTAARACAT | AGAGCTCTG   | CTGGGTACAG  | 14500 |
| TGCTGTCTCT  | TGCTTAATCT  | AGCAATTTGG  | GAGAGCTGCG  | CAGAGAGCAT  | 14600 |
| TTTCTGCTCT  | CACACACCTC  | CACACACCTC  | GCACACATCT  | CAGAGCCCCA  | 14650 |
| CACACACACA  | CACAAAAACA  | AGCTCTCTCT  | AGAAATGAG   | CTCAAAATTT  | 14750 |

CCCTCAGGTT CCTAGAAGAT CAGTCCCTCA ATTAGATTCA GATTGAGATG 14800  
 CTTTCTCTTT TAAACAATGA TTCCCTTTCT ATCATGCCCA ATAAAGAAAC 14850  
 AATAAAAAAT TAAACAATAC TGCCGTGAAT CTCAGCTACC CAGGAGGCAG 14900  
 AAGCAGAACT GCTTCAACCC GGCACACAGA AGTTGCAATG AAGTGAGACT 14950  
 GGCCTCATGC GCGCAAGCCT GGGAAACAGA GCAAGATTCT GTCTCAAAAA 15000  
 CAABACARS TCAATTCCTC CTCTAGTCC TGACACAGGA AATGTTAGAA 15050  
 ATATAGTCCA CAGAGAAAGA AGGAAGTAAG AATGTTTACG TGAATTTCT 15100  
 TGGAAAAAAT AGTTATACCT TCTTGCTTGT CTTCCTAACA GTCTCCAAA 15150  
 GCTTCGTACC TTGGCCAGAG GCTTGCTCTC TGCTGCACTG AGGTTTGGTG 15200  
 GCACCAAGAC AGACTTCCTA ATTTTCGATC CCAAGAAGGA ATCAACCTTT 15250  
 GARGAGAGAA GTTCTTGCCA ATCTCAAGTC AACCCAGGGT AAAATTTTTA 15300  
 AAGATTCACT ATATAGTCTA ATCTAGATCA GCTCTTGGGA GCTTAGAGCA 15350  
 GAGAAACTG TTTTCTTCTA CACCTAACCA TTATAGAGAT TAACTCTCT 15400  
 TCCCTCATC TGACCTGTGG AGGAATCTGA ACAAGGAGAG GAGGCAATGG 15450  
 CGAGGTTTCC TTATCATGAT GTTGTGATG TTCAGTGTGA GGCTCACAAA 15500  
 AAAAAAARAA AAAAAAARAA GGGCTCCTGG ATATAACTGA GAGCTCATTG 15550  
 TACAGTAAAT ATTAATAAAA CAGTATTGT AGCTGAAGAG TAGAACTGCT 15600  
 TGGAGGAGC AACTGGGTAG AATCGGCTCA AACTAAAGAG CATTTCTAGC 15650  
 CAAGACACCA ATGTATGATT GAGGATATT TATTCTAAAT ATGACATGT 15700  
 GGTGAACGAG ATCTGTGGAC TTCTGGGCTC CAACGTTAGA TTCTGATTTT 15750  
 AGCAAGCTTG TCAGGGGATT CTGATATTGA AAGGCTGTGG CTTCACTGTG 15800  
 AGAAACCTGC CCTAGGGGGC CATGAAATTT TGTCCTGTCT TTCAGAGTGT 15850  
 CTATCAGACCA TCAATGGGAA GTTAATCTGT ATCTTAACAA TACTTAGAGT 15900  
 GGGCCGCTGG ACTCACACT GATATCCCA CACTTGTGGGA GCTTAGAGCA 15950  
 GAGGATCACT TGAAGCCGAG GAGTTCCGGA CAGGCTGGG CAACATAGAG 16000  
 AGACGTTGTC TCTATTTTTT AATAATTTAA AGAGAAAAAA ATACTGAAA 16050  
 TATTGTATAC ACCACTGAAT TATAAATAG TGTATATAAT GTATATATT 16100  
 ATTATGAGGA ATATTTGAT ATTTCATATA TTATATCTTT TCCTTCTGTT 16150  
 TATTTTATCC AGTTATGAAG TATTTAGAAC AATTCAATG TAATTGGGGG 16200  
 AAAAAATGCA GAAAGTAAT CAGAAJAAT AGAAAAAGAG AGATGGGTTA 16250  
 CTTTGTGAATA GAGGCTGTT GTCTTTTGTG TTGTTTGGG 16300  
 GGGCTTTTTT TAGACAGAGT CCCACTCTGT TGCCGAGGCT GAGTGCAGT 16350  
 GGCACAGCA TGCCCACTG CATCCTTGAC CTCTTGGGCT CAGCAATCT 16400  
 TCCCACTTTA GCCTCCTGAG TAGCTGGGAC CACAGGTGCA TGTACACACA 16450  
 CCCAGCTAAT TTTTTTATT TTGTAGAGGA CAGTCTTCT ATGTATGCCA 16500  
 TCTGTATCTC AAACCTCTGC ACTCAAGTGA TCCCCTTGGC TTGGGCTCC 16550  
 AAGATATTGC GATTATAGCC CACGCCACCA CACCAACTC AGTTCATAT 16600  
 TAGACTTGGC CTTTCCCAAC CAGTCAATTG TGTCCAARAG ATCTCATAAA 16650  
 TGTAGACAGG AAACGTGCTC TTGCTCATCA GTTTTCTTCA TCTGTGTTCT 16700  
 AGGGGGATGG TCGGTGGGGG AAACCTGGGT TATGCAAGTT CTTCTGAAC 16750  
 ATCCTCTGTG AGCCACAGGA TGGATGAGG ACCAGCGGCC AGCGAGTCAG 16800  
 TCTGCAGCTT CACCAAGAGG AAGTCAATAG CTTCCAGCTC GCTCATAGT 16850  
 GCCAGCACCC GCCAACCTGT CTGTCTGTG ATAAAGAAAT GCTCTGCTG 16900  
 ACAGGATGGT GTGGATTTTT CTTTTTCTT TTTTTTTTT TTGAGACAG 16950  
 GTCTGGCTCT GTCGCCAGG CTGGAGTGCA ATGGCGGGAT CTGGCTCAC 17000  
 TGACGCTCT GCCTCCAGG CTCAGAGCAT CTCCCACTC CGGTCTCCG 17050  
 AGTAGCTGG ACCACAGGCA CACACACCA CGCCCAACTA AGTTTCTGTA 17100  
 TTTTTAGTGA AGCGAGGTT TACTATGTT CTCCAGGCTA GTCCTAACC 17150  
 CTTGAGCTCA AGCTATCCAT CTGCTTGGC CTCCCAAGA CTTGGAATTA 17200  
 CAAGCGTGAG CCACTGTGCC TGACACAGGT GGATTTTTTC AAGTGCARAT 17250  
 GTTGTGGTCC CAGAAGCTCT GATGTACCA AATTCACAG GAAAAAAGT 17300  
 CAATGGTACC CACCATCTCT ACCCTCCATG ATGGCAAGAG GAATACACA 17350  
 CACGTGCTGC AAGTGCCTAT TAAACAAAT TGCTATGGAT TTGAAGATG 17400  
 AAGTTTAGA GCTATGCTT AGTTTCTTT CATTAGATCT CTCTGTAAT 17450  
 TTCCAGCTTT TTTTTTTTTT TTTTATAGC AGTGTCTGCG TTTGTGCCCC 17500  
 AGTGTACCC AGGCTGGAGT GCAGTGACGT GATCTCGGCT CACTGCAACC 17550  
 TCCGCTCTGT GGGTTGAAGT GATTCTCTG CTTCAAGCTC CTGAGTAGCT 17600  
 GTATTTTAGT AGAGACGAGG TTTCACCAT TTGCCAGCT GGTCTCGAAG 17650  
 TCTGACCTG AAGTGATTGC CCGCATCTAG CTCTCCAAAG TGCTGGGAT 17700  
 ACAGGTTGTA GCACTGCACT GCGGCTGAT GTTCAAGAT TGTAGAGAG 17750  
 CCTTTGAATT TTTAATAAAC TTGTAGCTAT GTCCACATA TCCAGTTCA 17800  
 GTGTATGTC GATATTTCTT AGGAACACTG CCCTTGGTGT TTTTCTTTGT 17850  
 GGTAAATCAT GAGCCGCCAA ATTTGACAT TGTACAGAA TATACCTTTT 17900  
 CTCTGCTCT CTAACCTATA ACCAGAACTT AATTAATCTG TTTAGTCACT 17950  
 ATAAAGCTT AACTAATAAA TTTATTAGA TTTCAGCTG CTCACTGTGA 18000  
 AATAGGCTT TCAATGCTAT CTCTCACT TSCAGATAT TGCAAAATG 18050  
 GATCCATCCC TCTGATGTG GAGGAGAAGT TACGGTTGGA ATGCCCTAC 18100  
 CAGGAGCAAT TGCTACTCCG AGACACTAC CAGAAAAAGT TCAAGAACAG 18150  
 CACCTACTCA AGTAAGAAAT GAAAGGCACC CATAGAGTGT TCCAGCCCCA 18200  
 AAGATATTGT AATAGGTTGG ACTCGGCAC CAATCTAGCA AGTCTACGG 18250  
 AAGTTGATA AGCTGAAA TACTGABCA TTTCACAAAT GCGATCTCT 18300  
 AAACCTCAAA CTTCCTTTT GCTTTTTTGT TTTGTTTGT TTTCTTCAT 18350  
 CTGACATTG TTATGAGTCA CAGAATGAAA GATAAATCA TCAATCATGA 18400  
 TCTAACAATG ACCTTCAGTG CTCTAAAAA CACGAGGTG AAGGAAAAA 18450  
 TGAATATATT CTTATGTAA AATTAARATA CAGACATATA AAGGGCAAAA 18500  
 CATGACACTT ATTCATACCT TGAGGTTCCT CCCCTTCCA GAATAACCC 18550  
 CAGATCTCC TGGCTTGA GAGGGGCCCT GAGTACTCC 18600  
 AGACACTCTT GACCAACAG CAGCACTCTC TTTTTGTTCT CTTCTGCTCT 18650  
 TTTGCAAAAC CAGGCGTAGC TCAGCTACCC ATTAGTATGT TTTAGTCACT 18700  
 TAAACAGCT TCCAGTCTT CAAATTAGGA TGACATTGTC ACATGGGGCT 18750

18800  
18850  
18900  
18950  
19000  
19050  
19100  
19150  
19200  
19250  
19300  
19350  
19400  
19450  
19500  
19550  
19600  
19650  
19700  
19750  
19800  
19850  
19900  
19950  
20000  
20050  
20100  
20150  
20200  
20250  
20300  
20350  
20400  
20450  
20500  
20550  
20600  
20650  
20700  
20750  
20800  
20850  
20900  
20950  
21000  
21050  
21100  
21150  
21200  
21250  
21300  
21350  
21400  
21450  
21500  
21550  
21600  
21650  
21700  
21750  
21800  
21850  
21900  
21950  
22000  
22050  
22100  
22150  
22200  
22250  
22300  
22350  
22400  
22450  
22500  
22550  
22600  
22650  
22700  
22750

|             |             |             |             |             |       |
|-------------|-------------|-------------|-------------|-------------|-------|
| TTTCTTAAAC  | TACAGAAATG  | CAGGAAATGA  | TTACTCTTTA  | ACAAACAGT   | 22800 |
| ATGATGATAC  | ATGASAAATG  | GAGAGCAATTA | TATGATCAAT  | GATATCAAT   | 22850 |
| TTAAGACAAA  | CACATTAARAT | TATATTTTGT  | CCACTCTCAA  | AAAGTGTTAA  | 22900 |
| TAGACACAGT  | CACTGGTTTAA | AAGACAGAGT  | ACAGAACTGT  | CCAAACTTAT  | 22950 |
| GGCACTCTAA  | TCTGTCAGAA  | AACTTTTAAA  | ACGATCGCTA  | GGCCCAAAAA  | 23000 |
| TTGAGTCTAA  | TTGAGTCTAA  | TTGAGTCTAA  | TTGAGTCTAA  | TTGAGTCTAA  | 23050 |
| ATAATGGCTA  | ATAGTAAATG  | TGTCCTAAAT  | ATGTTGATAG  | TTTTTCTTCT  | 23100 |
| GCACAACTTA  | AAAGATCTCA  | ATGTCGCTCG  | TAAATTAAGT  | TGCGCTGGTA  | 23150 |
| GGAATGAAT   | TGACATAAAT  | ATATTTTTC   | TATATTAATC  | TTTTTCTTCT  | 23200 |
| CTTACTTTCT  | TTTTTTTGT   | TTTTTTTGT   | TGTTTGT     | TTTTGAGACA  | 23250 |
| GGGTCTCAT   | CTGTGCGCCA  | GGTGTGCTG   | TTGAGTCTAG  | TTGAGTCTAG  | 23300 |
| TTGAGTCTAG  | CTGTGCGCTG  | CTGTGCGCTG  | GATCTCTGCT  | CTCTCAGCTC  | 23350 |
| CCAAGTACAG  | AGCTGGGAGT  | ACGATGCTCT  | ACCATCATCT  | CGCGCTAAAT  | 23400 |
| TTGTGTGTTT  | TTGTAGAGAT  | GTGTTTGGC   | GATCTGTCCG  | AGSGCTGGCT  | 23450 |
| TGAATCTCTG  | GGCTCAAGCT  | ATCGCGCCCG  | CTCGGCTCCG  | CAAAAGTCTA  | 23500 |
| GAGATGACAG  | CATGAGCCAC  | TGCACTGTGC  | CGGCGGGCCA  | AGATTTATCT  | 23550 |
| ATAGTGTGTA  | TAGACATGAT  | TAAACATGAT  | TGAGTCTGTA  | CTGTCTGAGT  | 23600 |
| TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | 23650 |
| TACCCATGTA  | TTCACTGTGT  | CTCCTCCAGT  | TGCTCTTGCA  | AGAGTTTGGG  | 23700 |
| AACCAATAGT  | CTATAAATCT  | AGGCCAAAT   | TTTTATGTAA  | ATTTGTTTAT  | 23750 |
| TTTTTAAAT   | CTATAAATAT  | ACAGAGAAAT  | TTTTTAAAA   | TTTTTTTAAA  | 23800 |
| TATAATTAAT  | ATTATCAAAA  | TTTTTTTAAA  | TGCAACTGTG  | CATGAGTAGG  | 23850 |
| TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | 23900 |
| TCATGTGGA   | GCACATGAGT  | ATAAATTTG   | CTCTCCCAAC  | TTCTCTCGAT  | 23950 |
| GTCACTTGA   | CAAAACCTAA  | TGCTGACAGC  | AGAGCTTGTA  | GTATCTTCCA  | 24000 |
| TGGGATTTAA  | TTTCAACAGT  | GAGGACAAAT  | AAATGCAGAT  | GATTTTTTCT  | 24050 |
| AATTTGTCTA  | ATTTTGTGT   | CTCAAAACCA  | TAAATTAATC  | CATTATTATG  | 24100 |
| ACACATAGAT  | TTCTCAGTTT  | ACCAACAGAA  | ATAGTTATTA  | TTGAGTCTGTA | 24150 |
| TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | 24200 |
| CGTTTCTTCT  | TTAGGAAATG  | CTGTAGATTT  | GTCTATACAT  | TTTTGAAACT  | 24250 |
| CTCGAGGACT  | GGACTGTATC  | TGCTGCTTAA  | ATGCTGTTAT  | AGAACACAGCA | 24300 |
| GATTTGTGAT  | GGAACTGATC  | TAACTGCTAG  | TGCTCTCTGG  | ACTACGTCTG  | 24350 |
| TTCCAAGGGG  | TATAGACATT  | TTTGGGAAAT  | AGAGTCAATG  | GAGTACGCCA  | 24400 |
| GGGCAACAA   | CACTTAATAG  | GGAGTTCTCT  | ACAGGACATC  | ATCTCTCTCT  | 24450 |
| TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | 24500 |
| TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | 24550 |
| TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | 24600 |
| TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | 24650 |
| TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | 24700 |
| TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | 24750 |
| TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | 24800 |
| TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | 24850 |
| TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | 24900 |
| TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | 24950 |
| TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | 25000 |
| TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | 25050 |
| TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | 25100 |
| TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | 25150 |
| TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | 25200 |
| TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | 25250 |
| TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | 25300 |
| TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | 25350 |
| TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | 25400 |
| TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | 25450 |
| TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | 25500 |
| TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | 25550 |
| TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | 25600 |
| TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | 25650 |
| TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | 25700 |
| TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | 25750 |
| TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | 25800 |
| TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | 25850 |
| TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | 25900 |
| TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | 25950 |
| TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | 26000 |
| TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | 26050 |
| TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | 26100 |
| TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | 26150 |
| TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | 26200 |
| TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | 26250 |
| TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | 26300 |
| TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | 26350 |
| TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | 26400 |
| TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | 26450 |
| TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | 26500 |
| TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | 26550 |
| TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | 26600 |
| TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | 26650 |
| TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | 26700 |
| TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | TTGAGTCTGTA | 26750 |



|            |             |            |            |             |       |
|------------|-------------|------------|------------|-------------|-------|
| GTCTAGAT   | GCTGAGG     | TAGGACAT   | AGGATGAGA  | ATCATCTTAC  | 26800 |
| TTTCTTCTT  | TTCTTTTGG   | AGACAGATG  | TCACTCTCTG | AGCCAGCATG  | 26850 |
| GAGTGACCT  | GTACATCACT  | GCTCATCTGC | ATCTCTGCAC | TCCAGCGCTG  | 26900 |
| AAGCACTG   | CCCATCTCAG  | TGCCACAAT  | AGCTGGGACT | ACAGGTGTCAC | 26950 |
| TTCTCTCT   | CTGGCTACTT  | TGAAAAAAT  | TTCTTTGAGA | GTAGGGSTCT  | 27000 |
| TCACACTG   | CTCTCTCTG   | TGAAAAAAT  | TTCTTTGAGA | GTAGGGSTCT  | 27050 |
| CCCTACTCAG | CTCTCCAGG   | TGAGGAGATT | ACAGGACTAG | GCCACACACAC | 27100 |
| CCGACGACCA | TTCTTTCTTAA | ANAAAAAANA | AGATTCTCTC | TGTAGACAAA  | 27150 |
| TCCTCAATAG | CCATATCTGT  | ATTAAACAT  | CTGCTGCTCG | AATCATGTAT  | 27200 |
| TTCTCAAAA  | AGGAAATAT   | TGAGGAGATT | AGARATATGA | GGGATCTAGT  | 27250 |
| CGAAATGTC  | CTCATGATTA  | ATTATGCTAT | CAAAATGAGG | AGCTTTGTTG  | 27300 |
| AGGAGAGG   | CTCATGATTA  | ATTATGCTAT | CAAAATGAGG | AGCTTTGTTG  | 27350 |
| CTACGCTATA | TGTGCTGAT   | ATATGCTATA | TATCAGTAGT | TGATATAAAA  | 27400 |
| ATTTTTTTC  | TAGCTTCTG   | ATAGGCTGGT | GAGAAGTAGT | TGATTGACTT  | 27450 |
| ACATGGTCT  | AGTAAGTATG  | TCTCTCTAT  | TTAATACTAT | GARAATGAGG  | 27500 |
| TACGCTTAT  | TATTTTACCTA | TATCTCAAAA | AGTACTGCTA | TTTACTGCTT  | 27550 |
| AATGTACCT  | AGTTTAAATG  | TTTAAATGAT | TTTAAATGAT | TTTAAATGAT  | 27600 |
| AGATAGAT   | ATAAATATAT  | AAAAAATAT  | TTAGTTTAT  | TTAATATATG  | 27650 |
| TATGGTTT   | ATAAAGACT   | TTATCTATCA | GRAGACAGT  | GCCATACCT   | 27700 |
| CCCTGAGACA | TATAATGAT   | TAAACCAAT  | TGCTTGCCAA | AGAACACCTA  | 27750 |
| ACCCAGGACT | ACATGTCTCT  | GCAACTCTGT | TTTTTCAGA  | CAGAGATAGT  | 27800 |
| GATTCTTAGT | TACTAGGATA  | GAATGACTC  | CTCTCATCAT | CTCCCTCTCA  | 27850 |
| TTCTGGTCT  | TAAGGAGAT   | TAAGTCTAG  | CAATATACCT | AAAGACAGAA  | 27900 |
| AGTAAATTT  | GTTATTTT    | TCTGAGAT   | TTTCTTGTA  | ATTTAGACT   | 28000 |
| TTTGTAAAT  | GAGGAGCTG   | TACAGGGAA  | ATTTTCTCA  | ACCGTTAGT   | 28050 |
| ATTGGACAT  | TTTATTTCT   | CTGTGAGAA  | AGTTTTCAG  | GTAATAGTCT  | 28100 |
| TTTAAACT   | TTTAAATGAT  | AGCAGAAAT  | CTTATTTTAT | TTTAAATGAT  | 28150 |
| TTTAAATGAT | TTTAAATGAT  | AGCAGAAAT  | CTTATTTTAT | TTTAAATGAT  | 28200 |
| CAACACACT  | ATGACTTTAT  | CACTTTTAT  | TTTCCCTCT  | GCATTTGGCT  | 28250 |
| TTTCCOCAT  | TAGCTGATTA  | GAAATTTGAT | ATTATTCGAA | TAGCCTTTAT  | 28300 |
| TAGAAATTA  | TCACATAGAT  | ATAAGAAAT  | GTTGTGCTCA | CTTCTCATCT  | 28350 |
| CTCTCAGAC  | CTACTCTTT   | ATAAAGAAA  | AGTTTCCCAA | GCACCTAAAT  | 28400 |
| ACTATGAMCA | TATTTTATTA  | CTATATTTGA | AGGGTGTGAG | AGCTGATGCA  | 28450 |
| TTTCTCTCT  | TTTCTCTCT   | TTTCTCTCT  | TTTCTCTCT  | CCACAGCTCA  | 28500 |
| CTTTTCTTGC | ATTAATCCTAT | GATATCTTT  | TGTTCTTAT  | GTGGTGTAGA  | 28550 |
| GCACGACGG  | TGCGCARGAG  | TGCGCTTAG  | GAGAACAGCT | CTCTGCATAT  | 28600 |
| GAGGCGGAG  | TGCTCTTGT   | ATCGGACAG  | TTTTTCAGT  | GCTTTATGT   | 28650 |
| AGTGAAGCAG | CGCTGCTGAT  | AGGGTGCAGA | GTGACCTCT  | TCTCATCTCT  | 28700 |
| CTCTTTGCT  | GAATATGAT   | TTTCTCTCT  | TTTCTCTCT  | TTTCTCTCT   | 28750 |
| CTCTTTGCT  | GAATATGAT   | TTTCTCTCT  | TTTCTCTCT  | TTTCTCTCT   | 28800 |
| ATAAGACTG  | GAGGACGAT   | TTTACCAAT  | GTTTCAAAAT | AATCATCTA   | 28850 |
| AGGATGAAT  | TTTTCAGAAA  | TTTTGGCTT  | TAATTTATAT | CACTAAATAT  | 28900 |
| CTCAAGTAT  | CTCTCTCTA   | AGAGATGTA  | CTGTAAAGA  | ATGTCATATG  | 28950 |
| CGGATATAG  | TAGTATGAT   | TATATATAT  | TGCTGATG   | TGCTGATG    | 29000 |
| CCCTGCTGAT | TAGTATGAT   | CAACAGGTA  | AGACTGTCTT | TTTTTTTCCC  | 29050 |
| CTCTGCTCAC | CAGACTGAGG  | GCGACTGAG  | CGACTCACC  | TCACTGCAC   | 29100 |
| CTCTGCTCAC | CGAGTTTGA   | CGACTCTCT  | CGACTCAGG  | CTCAGTAGT   | 29200 |
| TGGACCTAC  | GGAGTTTGA   | CGACTCTG   | TAACTTTGT  | ATGATAGT    | 29250 |
| GAGCGGGGT  | CGAGCTCAG   | CTCTCTCTA  | TGCTGGGAT  | AGCGGACAT   | 29300 |
| CTCTCAGAC  | CGCGGACAT   | CTCTCTCTA  | AAAAAATG   | AAAAATGAA   | 29350 |
| TTAGAGATCA | TACAGCTTT   | CTCTCTGAG  | AGGATCATTA | TGTATGTATG  | 29400 |
| CTATATATTA | TAGATTCCCA  | AGAGATTAG  | AGCTTAAGT  | ATGATGCTCC  | 29450 |
| ACCGAGGGG  | TATATCATTA  | TTTAAAGAT  | TGCTGATG   | TTTAAAGAT   | 29500 |
| TATCTATTA  | ATGCTCTTTA  | ATCTATATG  | CTATATCTG  | ACATTAGAT   | 29600 |
| TTTTTTTCT  | TTCTCTTCCA  | TCTCTATCT  | TTTTCTCTC  | ATCTCATCT   | 29650 |
| TTTTTAAGCT | AGATATCACT  | CAAAATCTT  | TATGCTCTC  | GARGAAGAG   | 29700 |
| GATTAAGGAA | TGGAGATGTT  | TGTTTTCGCA | TTTAACTAAG | ATGTTGGGTT  | 29750 |
| TTTAAAGGAA | TGGAGATGTT  | TGTTTTCGCA | TTTAACTAAG | ATGTTGGGTT  | 29800 |
| AGCTGGTGT  | CAATTTTAT   | TTTATTTACT | TTTCTCTCT  | CTCTGTTAT   | 29850 |
| TTTTTCTCT  | AGCCCTTCA   | AGTTTGTGT  | GAGGAAACT  | CATAGGACCA  | 29900 |
| GGAATTTGCT | ATAGAGCTGG  | GAAATTTGCT | CACCTTCAAT | GGGATACCT   | 29950 |
| TACATGATTA | ATAGTCTCT   | AGGATTTTCT | ATCATATGGA | AGATATGAAA  | 30000 |
| GGGAGGAGTA | CTGCTGCTCT  | TGCTGCTGCT | TTTCTGCTG  | CTCTGTTAT   | 30050 |
| CTCTGCTCT  | AGATCAATTA  | ACGAG      | GGCTGTGGT  | CTGACATCT   | 30100 |
| TGCTCTGTA  | AGACCCACAC  | AGGTTCACTT | TGCTGTGCT  | TACATAGATG  | 30200 |
| CCATATTTA  | AGACAGAGGCA | ATTGCAAGG  | AGAAAGAGTA | ATTATAGCAG  | 30300 |
| AGCAGGCTGT | CGAGGAGGCA  | AGAGTTTAT  | TATTACTACA | AGTTTGGGCT  | 30400 |
| TTTAAAGGTA | CGAGGAGCTG  | AGGTTTAT   | TTTCTGCTG  | AGTTTGGGCT  | 30500 |
| TTTGAAGGTT | GAGAGTCTGT  | GTTGCTCAG  | TGTTGAGATG | AATCAGAGG   | 30550 |
| AGTGAAGGTA | AGGTTTCTCT  | CTGTCTCTT  | GTTCTCCTAT | GGGATGGCAG  | 30600 |
| AACTGGTTGG | GCGAGATTA   | CGGTCTGGT  | GTTCTCAAT  | ATTCCACCCA  | 30650 |
| TGTAGGCTGT | TGCAAGATAT  | CTAGCAAGCT | TAATCTAGT  | TTTCCACAG   | 30700 |
| TGATGTTATC | CCGAGAGATA  | GGTATCTCT  | AAGTGTGTA  | CTAATTTGTT  | 30750 |
| AGTCTCTCA  | AGGATAGACT  | TCTCCGAGG  | AGAGAGGGG  | TTTTTTCAGA  | 30800 |

AAGGGCTAT TATCATTTTT GTTTCAGAGT CAAACCATGA ACTGAATTC 30800  
 TTCCCAAGT TACTGTGACC TACACCAGG AATGAGAGG GACAGCTTAA 30850  
 AGGTTAGAG CAAGATGGAG TCAATGAGGT CTGATCTCTT TCAGCTGTCA 30900  
 AATTTCTCTA GTTATAATTT TTGCAAAAGG GGTTCAGTC CCAGCTACTT 30950  
 GGGAGGCTGA GACAGGAGGA TTAATGGAG CCAGGAGTTT GAGGTTGCAG 31000  
 AGAGCTAGA TCACGCCACT GCACCTCAGC CTGGGTGACA GAGTGAGAC 31050  
 CTGTCTCTAA ATAAATAAAT AGAATAAAT ATAAATAAAT AAATAAATC 31100  
 AAGTGGTGT GCAATAGAA TTGAGCGAT TTGTTCCAA ACCTCAAGAA 31150  
 AGCTTGGTCT TGCTCTGTCC CAGGTGGCTG GATAAATGG GCCTGTGACG 31200  
 CGRATGGGA ATAGAAAGTG TGATGAGGCA AGTATCTTT GGAGCAGGAA 31250  
 ACTACCATT AGTGATGAG AACTTCGATC TTTACTCTGT AAGTGACAT 31300  
 TATTTTCTTA ATTCAGTGG AGTGATTTAA AGTCAACTCA GGACCTCTGG 31350  
 TGTATCACTC CTATGAACAG TCAGTCTCTT CAGTAACTAG CCAATCATG 31400  
 AGTGTAGTAA CTCTAGAGG CTCTAGAGT CATTCAATCT AACATTTTTT 31450  
 TGTGTGTTTG AAGAAGAGAA ATCAAGAGCT AGGAATACT TTTTAAAGT 31500  
 AAGCCATTGG CAGTATAGTG TGGATTTTGT TAAAAGGGG ATAATTGAA 31550  
 ATTTTATGAC TCATTATACA AGACAAAATA AGTTGGATT TCAATTTGT 31600  
 TACAAAGTAA ATCAAAAGTA TAATTTGCTA CAGTACGCAA AGCTTCABAA 31650  
 CATTTTATAT GTATAGAAAT TGTAATTTAT TTAACTTAA AATGAGCCAG 31700  
 TACATGTGT TTGCTTAAAA ATCTCATGCT AGAANTTTC TATGTTGTTA 31750  
 ATAACTCTCA AGATATTTAT GAATAAGTTC TTAATTTCTAA TCCTTCTCTC 31800  
 AACTGTATCT GGTGCTAAAT CAGGAATATG TTCTTCCCAA AAAGCCTCGT 31850  
 GGAAGATCTG TATGCTTAAA TATATGTGAG GGATAATACA GATGTAGGCC 31900  
 TCGAAGCAT GACCTTGATT TTTATAGTCT AAAATGTCAT TTGCAAGTAT 31950  
 CTATTTTCTA AGAATAATTC CTAAAGAAAT TATTGAAAT TTGTAGGAA 32000  
 GCTAAGAAAT TTTCAGAGCA GGTCACTGGA AAATAAAGC TAGGCTTTTG 32050  
 TGGTTTGTGG ATGACCTTCC CAACAAAAT GCTTTTATCT TATAGTATC 32100  
 CAAGCTTGTG GAACATATTA GTCACTCTTT TTTAGAAAT TCTTAGAAAA 32150  
 GTGATCTTGC AAAAATGGAA TTTATCTTTC CCAAGATATA TCTGTCTAG 32200  
 TATAGAGTTA AACTAAGCAT AGTAATTTCA CCAGCAACAC ATTCAAAATC 32250  
 TACTCTGACG CTTTTTATCT CATCCAATTT TTCCAGGGG CAGACATTA 32300  
 ACCTTGCTCT TGTGTATGCT TGTGTATGCT ACTAAATGGS CTCTCTCTT 32350  
 AAGGTTCTCA GTCACTAGA AAATGTGCA AGAGTAATGT GTACCTCTCT 32400  
 CACTTGTAGA TCCAAGAGAA TTAGACTTAA ACTCACTA CATGTCTGTG 32450  
 ACTTTTATTT ATTTGCATGA CAGCTCTGTG AGGTGGCAG CGAGTATCT 32500  
 TGGATCCATT TTTTAGATTA AAGAGTTCAA ATTAGAGAGA GGTTCGTGA 32550  
 TTTCAGAGAA GCCATACGT AGTCTATGT TACTCTTAAA ATTCGATTC 32600  
 AAATCTGCT GTTCTGAGCT GATCACTCTA TACTCTTAAA TCTTGTACC 32650  
 CATGCTATG TCTCCTTCTA AAACATTGAT TCCACTCTTG TCTCAGTGA 32700  
 AAAAGTGGAA TTTAAGCAGA GAACAAAGAG CCAATTTGCT TGTAACTCT 32750  
 ACTTTCCCTC TACTTTCAAG AAGGAAAGTT GGGGTATGTG TTGAATGGT 32800  
 ATTTATTTAT TTATTTATTA TTTTAAAAAT TGATCAAGAG TCTTACTGA 32850  
 TTGTGAGGCG TGGTCTCAAA CTCTGGGCT CAGTGATCA TCACACTCA 32900  
 GCTCCCAAGT GCTGGGATTA CAGCATATCT CATTTGGCC ACCACGCTG 32950  
 CGCAGTTTTT TAAGAAAAAC TTTTACTATA GAAAATTTTA ATCATATACA 33000  
 AAATACAGAG GAAAGTATAT GAACCCACTT TAGGAGACTA GAATATGCCA 33050  
 CCCCAAAATA TGCCACTTTG GCATAAGGAT TATTTGAGC TAAAGGCAC 33100  
 TGGGAGAA GACATAGAG AAAGGTTCTC TGCTCTTCTC CATTGTCTTA 33150  
 AAGCAGGAG ATGAATCTTA AAAGTCCGCC TCCCTCCCT TCTACAGGA 33200  
 AAACAGAG TTATCACTG AGNTAACTT CAGACCCCTA TCAAGTGTAGA 33250  
 GATGCGACTA GAAGAATCTA TATTACATAC TCAATTTATT TCCTTCCAC 33300  
 AACTTGGCAC CCCAGAGACT AAAAATCCTT TTCTTTTGT ATGCTCTTGG 33350  
 TCCAAAATTT TGCTCTATAA GCTGGAGTTC TAAGCCACTT CTITGAGAA 33400  
 TACTTGTCTC CTGTGATTTT CTGTTAAMAT ACATGATTA ATATACATGT 33450  
 TAACAAGCTT TGTGTGATTA TTCTCTGTT TGTGTCTGA TCTAGAGGT 33500  
 CCATCCAC TAAGACTTAA AGAGTAGGAG GAAATATATA TTTCTCTGT 33550  
 CATACTTTGA TCTTGTTTAA TCCGTARCCC TTCCCACTTT TCACTCTCTA 33600  
 CCTATTAGAT TACTTTGAAG CAAATTTTCA ATATATTACT TTATCTATAA 33650  
 ATATTTCAAT ATGTGCTAGG TGTGTTGGCT CACACTCTGA ATCCACAC 33700  
 TTGCGAGAG TGAGGACAGA GBRATCTTGT AGCCACAGAG TCTCAGACA 33750  
 GCTAGGCCAA AAAAATCTTA AAACCTTAT TGCCAGTGT GGCACTGCC 33800  
 TGTGTCOCA GCTACAGAG AGGCTGAGGC AGGAGATTCG CTTAGGCCA 33850  
 GAGGTTGAG GCTGAGTAA GCTGCATTCA CACCACCTGA CTCAGCCTG 33900  
 GGTGACRAG TAAGCCCATG TCTCAAAAAA ATACATATTT TAGTATGAT 33950  
 CCTTTTGTGA AAAACACAT ACTTTTATTA TACTTTAAT AMTAACTA 34000  
 ATTCCTTAGT ATCCAAATAT ATTTTGTGCT TGCTCTCAGT TTTCTTAT 34050  
 GCTARAATA TTGTTATAGA TTTTATCAAT CAGATGCCAA ACAAGTGTG 34100  
 TATATTACT TTGTTGACA AGTCTTTTAA GTTTTTCAT CTTTAACTC 34150  
 TTCTCCCTCT TCTTTCTCT CTGTAAATTT ATTAATGTGA AAAACAGGT 34200  
 AATTTGTCT ATAGTATTC CTATAGCTAT TCTCTGTGCT CCTGTGTTT CTTGTAATC 34250  
 CTATGATATC ATTAGAGTCT TTCCCTGTGCT CCTGTGTTT CTTGTAATC 34300  
 TGAGTATATA CTTAGAGCT TGAGTTTAT TTGCTTTAT ATGTTAT 34350  
 TTTTGCAGA ATCTGTATAT ATCTGCTCT GAGGACAG AATGCTGTG 34400  
 TGCTCTGGT TTGATCTTG ACAGCTACTG ATGACCATG CCTAATCAT 34450  
 TACTTTATG GGTGGGGGG AATAAGGTTT TAAATAAAT TTTTTTAA 34500  
 GATTTTITA ACTGTATTT TGAGCAGCTG TCTCATTTG TTTCCAGGC 34550  
 TGGAGTGCAG TGGCACAATC AGCGCTGAG TCACCTTGA CTTCTGGGA 34600  
 TCAGGTGATC TTTCTACTC AGCTCTCTGG GTCTCTGCT CTACAGTGC 34650  
 ACACACAC ACCTGCTAA TTTTGTGAT TTTGTGTA GAAGGGGTT 34700  
 CATCATGTT CCAGACTGG TCTTGAACCT CTGGGTTCAA GTTGTATCC 34750

CACTTCAGCT TCCCAAAATC CTGGGATTAC ACTTTGGGCA CGTGGCGTGG 34800  
 CTTAAATGAA ATTTTCTTGT TCTTAACAGA CAGAGCTTTT ACTTTAAAAA 34850  
 TTGTCTCTTG CTGTGATCAT TGTTTGTGTA TGTGTGTGTG TCTTAAAGTT 34900  
 TGCGTTTGAAG CTITGCTTGG AATTCTTTGGA TGAACAATAA CCAAGAATAC 34950  
 TTAACATCTG ATCATTTCTG ACAGATATCC CTTACAGGCT ATGGCTTTT 35000  
 GAATTTGTGTC CTCCAGTGTG AAAAAGCAGC AAGCACGATA CTGCTCTCAG 35050  
 ATTCATGGTG GTACATATGT AGTGTAAAAA AAAAARAAG ATGATACCTA 35100  
 TTTAAATGCC CCCAGATATA CAGTATGACT CTGTGAGGA TACTATTGTT 35150  
 CTGTGCACTG GTTTCATATA ATBAGGACAT AAGTAARGAT CTATTTTGT 35200  
 CTTCTTTCTC CCAACCACCA CAACTAGGAT TATTGGCTAT CTCTCTGT 35250  
 CAAGAAATGT GTGGGACACA AGGTGTTAAT GGCAAGCGTG CAAGTTCAA 35300  
 AGAGAGAGGA GCTTCGAGTA TACCTTCATT GCACAACAC TGACAGATTA 35350  
 GTATGAAMCA CACCCTTTAT CATATCATCA GTTTTAGTGG GTAGTACCTA 35400  
 AACCTTTACT AAACCACTTG TGTGATGTGT CTATCATTC CATAGATATA 35450  
 GCGACTGTCA ATTAAGTAA TTTTATACA ACGAATTTAT TTTATTAT 35500  
 TTTTAGAAGA AAAATGCTAT TTTTGTGTGT GTTGTTTTTT GAGCAGGGGG 35550  
 CTGGCTCGTC ACCCAGGCTG GAGTGCAGTG GTGCATCTCT AGCTCACTGC 35600  
 AACCTCCGGC TCCCGGGTTC AAGTGATTCT TGAAGAGGAG AACRAATAA 35650  
 ACAACATAT TAITTTTCAA AGTGTGACCC GCAGTTTCTG GAGTTAGAAA 35700  
 GACATCGAGA TTTTGTAGC CTCACTCTCT TGTCTTAGGT AGCAAAAT 35750  
 TGTATGATG ATGCTGAAAT ACTRAATTCTA GCCAAAAAG ACCAGCTACC 35800  
 ATTTCCGATT GTTGGGAGCT GGGAACTCTG GATAGTAGG ACCCCAGTAG 35900  
 GAAGTAGGCA GGGGAATGTT TGAATGGAT AAATTCATTA AAAATGTAG 35950  
 TAGATTTAT TTTCTTATAC ATTTCACTCT TTTTATAGG CTAGAAAAAG 36000  
 CCCCCTTTT TATGGTTTTAT TTTTGAATG ACTTCAAGC CACAATTT 36050  
 GCTTTTACC TCTGATGTC TGAATATGGA TAGTCTGGCT GGCTCTTAA 36100  
 CAACCCAGCT GGCAGAGCTG TGAGGATCTC AGTGTGCTCT AGCCGACAGA 36150  
 TTGGTAGCAT GRAACGGCAAC ATTTTAAAT GTGTTTTCAA AATGAGACA 36200  
 CACTACGGGT CTAAAAAGAT CATAAAAGAA GGATACATAG AGGCGCCACT 36250  
 GTCATATGG ATCTCTATAC TTAGAGTGA TATGAGTGT TCAATATGA 36300  
 TACTAATACT TAGGATCACA TTTGTAATG AGTTTAAAT TGTCTAAAT 36350  
 AGATACATAT TGAACCTCTG TATGCCATAA ACCCTCATTA TGFCAACAG 36400  
 AAGAGAGAGA TTTAACTCTG TATGCCATAA ACCCTCATTA TGFCAACAG 36450  
 TACTTCCGGT TACCTTATCC TTTTCTAAC AAGCAAGTGG ATAATACTCT 36500  
 TCTAAGACCT TTGGGACCTC ATGGATTACT TCCAAGTAA GTAATTTTCC 36550  
 TTGTTCAAT CAAACTTTCA ATAAATTTAT TGGTGTATT CAGAATAGAG 36600  
 AGTTTGAGCA GGGAGCAAAA CACAAGTCA ACTATATAGA GTCTCAATTA 36650  
 TCTTAATAT TCTGAGAAAT TATGATGTA TACTACTATA TATGATATA 36700  
 ACTCATCTTA AGAGTCTAAA GCAAAAGAT GTGACACAAA ACTAGAGATT 36750  
 ATCTTAGAGA ATAGTTTTCG ATTTCAAAAT AACTTGACAT ATCAAGATCC 36800  
 ACTCACAGCA TTTAAATTAI TTACTTAAA AGACATATAT TCTTGGTAC 36850  
 ACATTCACCT AAGCAAAATA TACCTTTATA TAATGTGAT CAAGGTATG 36900  
 TGGGTTGTA TAAATATCTA TACCATGTA GTACAGTGT ATTCCTTAT 36950  
 AGCATTAAT TTTATGCTT AGAGTAGAAA AAGAGATAGC TAGAGTATAT 37000  
 TTCTTAAGTA GATTTCTATA CACTTTGGTT TCAAAAACCA ATTTATGACT 37050  
 ACATCTTATA AAAGCTGTGA TCAATGGAG TGCCAAAAAA TGACTATGAG 37100  
 TCTTAAAGAG TTAGGCATAT AAATATTTA AGTTTTCTGT TCAATGTATG 37150  
 TTGGAAGAGG TTCTTTTCTC ATGACTATTC TCATATTGGA GCATAAAG 37200  
 AGTTTACAGC CTGGCCGAGC TGCTCATGTC CTGTATGCC ATACTTTGG 37250  
 GAACCTGAAC CAGCAGATCT ACTTCAGCCC AGGAGTTTGA CAGCAGCTGT 37300  
 GGCAATATGG CAAATCTCTC TCTACAAAT ATACCRAAT TAGCCAGGCG 37350  
 TGGTGGTGCA TGCTGTAGT CCAAGTACT TGGGAAGCTG AGGTGGGAGG 37400  
 ATFGCTTGAAG CCCAGGGGGG TCATGGCTGC AGTGAGCTGT GATGTGCTCT 37450  
 CTCTACACCA GCTCGGTGGA CAGAGTAGAG CCGTGTCTCA AAARATAA 37500  
 TAAATAAATA TTAGAGCTTT ACMAAATCT CACATCTC TCCACTCTT 37550  
 CCAATPGCA CATAGTAGAT GTGTCCAGS ACTATTAGCC TCGGAGCTGT 37600  
 AGGCAGTACA GTAAGCAGCG TTTCTCAA TGTCTGTCCC CCACAGACAA 37650  
 ACATATTATA CACTGGGTAC TGCTTTTTTA TTTTTCCTC CTATGCTTT 37700  
 ATTTTACTAT AACTATAAT ATATACATG TAATAGBAA AAGCAGGCT 37750  
 CGGGGGAGAG ATCCAGAGAT TCTCCCAAG GCTTCTCA ATGAGCTCT 37800  
 CTGAGCATTT TTGTTCTCT CTCTTTTCT TTTTTTTT TCTGAGACA 37850  
 GAGTCTACAT CTGTTGTCCA GCGTAGAGTC CAGTGGCGTG ATCTAGCTC 37900  
 ACTGCAACCT CCGCCTCTG GGTTCAGACA ATTTCCCAAC CTCAGCTCTC 37950  
 CTAGTAGCTG GGATTAGAGG CATGCATCAC CAGCGCTGGC TAATTTTGT 38000  
 ATTTTGTAGT GAGTAGAGGT TACCAACTGT GGGCAGGCT GTCTGTGAC 38050  
 TCTTGACCTC AGCTGATCCA CCGCTCTTAC CCGTCCAAAG TGCTAGAGT 38100  
 ACAGCAGTGA GCGAGCTGCG CTGCGCCTA TTACATTCTG ATCNCACTT 38150  
 TCACTTTTAT TAATTGCAAA ACTGGTGAA TTAATAGCAA TGTTTGTCT 38200  
 CCTARATTC TCTTTGATGA GTATATATTA CTACACTCT TCTGTCTTA 38250  
 AATTTTGA CAATATGATC CTAGATAAGT TTAAGTGTG ACAGCTGTA 38300  
 GGCTTACTCA TATTAAGAC CTGGAGAGT TAACACAGC TCACCTTTAA 38350  
 AAATATATAC TATCATATCT ATTAATTTG AGCGGGGGT CTATCTG 38400  
 CTCCAGGCT GAGTAGATGT GTTGGGTCA CAGCTCACTG CAGCACCGC 38450  
 TACCTGGCT CAGTGTATCT TTCTCTCA GCTTCTGAG TAGCTGAGC 38500  
 CACAGGCTTA TGCTACCACA CTGGCTTAT TTTTAACTT TTTGTAGAGA 38550  
 CGATGTCTCA TTATGTTGCC CAGGCTGTCT TCAAACTCT AAGCTCAAGT 38600  
 ATCTTCTCTC AGCTGCCCAA AGTGCTGGGA TTACAGGCT GAARACAT 38650  
 GAGCCGCCCT AAAAATATT AGGCTCTGC ATAGTAGAC TTTATTAAT 38700  
 ATTTAAATGA ACATCTGGTT TTTTAAAAA AAAATAGAG ACAAGCTCTC 38750



CTCTGTGCAG TGTGTAGT CAGCGAAGA CTATATACC TGTGGGGACA 42800  
 GCGATTAGCC ACCACRACCA GTCTTATTTT AAGTTTATTA AAAATGGCTG 42850  
 GGGCAGGTGG CTCACGACCTG TAATCTTAGC ACTTTGGGAG GCGAGGCGAG 42900  
 ATGGATCACC TGACCTGAGG AATTTGAGAC CAGCTGGGCG ACATGTGTGA 42950  
 AACCCCATCT CTACTTAAAAA ATACAAAAAT TAGCTGGGTG TGGTCTTGTA 43000  
 GTCCACAGCTA CTTGGGAGGC TGGGCGAGGA GAATTAAGTG AACCCAGGAG 43050  
 GCAGAGGTTG CAGTGAGCCG AGATTGTGCC ACTGCACTCC AGCCTGGGTG 43100  
 ACAGAGAGAG ATTCTCATCT AAAAAACAAA GTTATTAAAA ATGTATATGA 43150  
 ATGCTCTTAA TATGTCTCAG AAGCAGGAAA GCGAGGGATA TATTATGAGT 43200  
 TTTTAGAGAG TGTCTAGCTG TATTTATTC TTTCAAAATG TTTTAGAGGA 43250  
 TTTTAGAATT CTTTCTCTCA TGTGCCATCT CTACAGGCAC CCATCAGAAA 43300  
 AAGCATACTG CCGTACCGGT GAACCTGGTT GTAAAAAGAA AACTATCTAT 43350  
 TTGCAACCTA AAAGACAGCT AGATTTTGCT GATTTTCTTC TTTGGGTTT 43400  
 CTTTGTGAGC AATATATATG GAGAGGACAG ATTGTAGAT ATGATAGTAT 43450  
 AAAAATGGT TATGACCTC TACAGAGGGA GCGAGGGATA TAACTTAAA 43500  
 ATTACTATTA ATGAATATGA TTTGTCAAGA GGATAAAATT TAGAAAACAC 43550  
 CCAATACCTT ATAACTGTCT GTTAATGCTT GCTTTTCTC TACCTTTCTT 43600  
 CCTTGTITCA GTTGGGAAGC TTTTGGCTGC AAGTAACAGA AACTCCTAAT 43650  
 TCAAAATGGC TAGCAGATAA GGAATGTAT ATTCCACAT AACTAGACGT 43700  
 TCAAAACAGC CAGGCTCCAG CACTTCAGTA CGTCAACAGG GATCTGGGTT 43750  
 CTTCCACGCT CTCGCTCTG CACTCTTAG CCGTGGCTTC ATTCTAGAAC 43800  
 TCTGTAGCA TGATGGCTGT AGCTGTITCA TGGGCCCTT CAACCTCAT 43850  
 AGCAACCGAG GGAAGAAAT GAGCAATTT TTGAGTCTCC TTCATAGACT 43900  
 TGAATAACTC TTTTTCAGAG CTCTCCACAG CAAACCTCTC CTCATGTCTC 43950  
 CTCATGTCTT ATTTGTCAGA AATGGGTAAT GTGGCCATTT CACCAATCAC 44000  
 TGCCACACAC ACGAGGTTTC CTATAATGT CTCAGATGA CCCCTGGAAA 44050  
 TGGGAGGGGT GTTGTGAGCT CTACAAGTC AAGCATCGAC TTCTCGGCTT 44100  
 TTTACCACTG AAAAAATGTA ATTAATTTTC CCTCTAAGG ATTAATATTC 44150  
 TTCAAAATGA TGCCTGTTAT GGATATAGTA TCTTTAAAA TTTTATTTT 44200  
 AATAGCTTTA GGGGTACACA CTTTTTGCTT ACAGGGGTGA ATTGTGATG 44250  
 GGTGAAGACT CGGCTTTTAA TGACTTTGTC ACCTGAATGA TGATCATGTG 44300  
 ACCCAATAGT TATATTTCAT TCAATTACC TCTCTCGCC CTCTCCCTT 44350  
 CTGAGTCTCC AACATCCCTT ATACACCTGT GTATGTTCTT GTGACCTAC 44400  
 AGCTAAGCTT CCACTTATAA GTGAGACAT CGAGTATTG GTTTTTCATT 44450  
 CCTGAGTTAC TTCCCTTAGG ATACAGGCC CCTGTTCCGT CCAAGTTGCT 44500  
 GCARAATACA TTATCTTCTT TTATGGCTGA GTAATAGTCC ATGGTACATA 44550  
 TATACCATAT TTTCTTATCT CACTATCAG TTGATGGACA CTGAGTTTAA 44600  
 TTCCATCTCA TTTATTTCTA TTTAGTATA TTTTAGAGA GTTAAAGTGA 44650  
 AAAATTAAT TTTAGATCTT TCAATACCTT TAAATTTAT ATGTAAGTGG 44700  
 TTTTATATTT TTCACTATTG AAATAAAGTA ATTTTATATA CCTGATATT 44750  
 GTATGACTAT TCTTTTAGTA ATGTAAGGCC TACAGACTCC TACATTTTGA 44800  
 ACCACTAGTG TGTGTTTCA CCCCTTGTA TACTATCAGG ATCTCTGA 44848

## (2) INFORMATION FOR SEQ ID NO:43:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2396  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (X1) SEQUENCE DESCRIPTION: SEQ ID NO:43

TTTCTAGTTG CTTTTAGCCA ATGTGCGATC AGGTTTTC AAGCAAAAG 50  
 AGATACTGAG ATCTTGGGCA GAGGACATCC TAGCTCGCT AGATTGGGC 100  
 AGGCTCAAGT GACCAATGTC TTAGGCGAGA AGGAGTCCG GGTAGGCTCT 150  
 GGCTGAACCC TCAACCGGGG CTTTAACTC AGGGTCTAGT CCGTGGCCCA 200  
 AATGGATGGG ACCTAGAAAA GGTGACAGAG TGGCAGGAC ACCAGGAAGC 250  
 TGGTCCACCC CCGTGGCGGC TCCCGGCGCG TCCTCCCCA GGCCTCCAG 300  
 GATCTTGGAT TCTGGCCACC TCCGCAACCT TTGATGGGT GTGGATGATT 350  
 TCARAAATGG ACGTGACCGC GCGCGAGGGG AAGCCAGCA CGGAATGAA 400  
 AGAGAGCGAG GAGGGGAGGG CGGGAGGGG AGGGCGCTAG GAGGGGACTC 450  
 CCGGAGAGGG TGGGAGGAT GAGCGCTGT GGGAGGGTAC TGAGTCTG 500  
 GCACAGAGGC GAAGCAGGAC CGGTTGAGAG GGGCTTAGC CAGCGGCGCG 550  
 GCTGCCCCAG CTCTCCCGGC AGCGGGCGGT CAGCGAGGT GGGATGCTGA 600  
 GGCTGCTGCT GCTGTGGCTC TGGGGGCGCG TCGGTGCCCT GGCCAGGGC 650  
 GCCCCGCGCG GGACCGCGCC GACCGACGAC GTGGTAGAT TGGAGTTTAA 700  
 CACCAAGCGG CCGCTCGGAA GCGTGAGTCC CTGTTCTCT TCCATCACA 750  
 TCGACGCCAG CCGTGGCCACC GACCGGCGCT TCCTACCTT CCGTGGGCTCT 800  
 CCAAGGCTCC GTGCTCTGGC TAGAGGCTTA TCTCTGCAT ACTTGAGATT 850  
 TGGCGGCACA AAGACTGACT TCTTATTTT TGATCCGGAC AAGGAACCGA 900  
 CTTCCGAAGA AAGAAGTTAC TGGAAATCTC AAGTCAACCA TGATATTTG 950

AGGCTCTGAGC CGGCTCTGCG TCGCGTGTTG AGGAACTCC AGGTGGAATG 1000  
 GCCCTTCCAG GAGCTGTTGC TGCTCCGAGA GCGTACCAAA AAGGAGTTCA 1050  
 AGAACAGCAGC CTACTCAAGA AGCTCAGTGG ACATGCTCTA CAGTTTGGCC 1100  
 AAGTGCTCGG GGTAGACCT GATCTTTGGT CTAAATGCGT TACTACGAAC 1150  
 CCCAGACTTA CGGTGGAACA GcTCCAACGC CCAGCTTCTC CTTGACTACT 1200  
 GCTCTTCCAA GGGTTATAAC ATCTCTCGGG AACTGGGCAA TGAGCCCAAC 1250  
 AGTTTCTGGA AGAAGCTCA CATCTCTATC GATGGGTTC AGTTAGGAGA 1300  
 AGACTTTGTG GAGTTGCATA AACTTCTACA AAGGTCAGCT TTCCAAATG 1350  
 CAAACTCTA TGCTCTGAC ATCGGTCAGC CTCGAGGGA GACAGTTAAA 1400  
 CTGCTGAGGA GTTCTCTGAA GCGTGCGGA GAAGTGATG ACTCTCTTAC 1450  
 ATGGCATCAC TATTACTTGA ATGGAGCAT ATCGTACCAA GAAGATTTTC 1500  
 TGAGCTCTGA TGCGCTGGAC ACTTTTATTC TCTCTGTGCA AAAAATTCG 1550  
 AAGGCTACTA AAGAGATCAC ACCTGGCAAG AAGGTCGTGT TGGGAGAGAC 1600  
 GAGCTCAGCT TACGGTGGGG GTGCACCCCT GCTGTCCAAC ACCTTTGCAG 1650  
 CTGGCTTTAT GTGGCTGGAT AAATGGGGCC TGTGAGGCCA GATGGGCATA 1700  
 GAAGTCGTGA TGAGGCAGGT GTTCTTCGGA GCAGGCAACT ACCACTTAGT 1750  
 GGATGAAAC TTTGAGCCTT TACCTGATTA CTGGCTCTCT CTTCGTGTCA 1800  
 AGAACTGGT AGGTCCCAAG GTGTACTGT CAAGAGTGAA AGGCCAGAC 1850  
 AGGAGCAAC TCCGAGTGTA TCTCCACTGC ACTAACGTCT ATCACCCACG 1900  
 ATATCAGGAA GGAGATCTAA CTCTGTATGT CCTGAACCTC CATATGTCA 1950  
 CCAAGCACTT GAAGTACCG CCTCGTTGT TCAGGAAAC AGTGATACG 2000  
 TACCTTCTGA AGCCTTCGGG GCGGATGGA TTACTTTCCA AATCTGTCCA 2050  
 ACTGAACGGT CAAATCTTGA AGATGTTGGA TGAGCAGACC CTGGCAGCTT 2100  
 TGACAGAAA ACCTCTCCCC GCGGGAAGTG CACTAAGCTT GCTTGCCTTT 2150  
 TCCTATGTT TTTTGTGAT AAGAAATGCC AAAATCGCTG CTGTATATG 2200  
 AAAATAAAG GCATACGTA CCCCTGAGAC AAAAGCCGAG GGGGGTGTTA 2250  
 TTCATAAAC AAAACCCTAG TTTAGAGGC CACCTCCTTG CCGGTTTCCA 2300  
 GAGCTTCGGG AGGTTGGGT ACACCTCAGT ATTACATCA GTGTGGTGTG 2350  
 CTCTCTAAGA AGAATCTGC AGGTGGTGAC AGTTAATAG ACTGTG 2396

## (2) INFORMATION FOR SEQ ID NO:44:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 535  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:44

Met Leu Arg Leu Leu Leu Trp Leu Gly Pro Leu Gly Ala  
 5 10 15  
 Leu Ala Gln Gly Ala Pro Ala Gly Thr Ala Pro Thr Asp Asp Val  
 20 25 30  
 Val Asp Leu Glu Phe Tyr Thr Lys Arg Pro Leu Arg Ser Val Ser  
 35 40 45  
 Pro Ser Phe Leu Ser Ile Thr Ile Asp Ala Ser Leu Ala Thr Asp  
 50 55 60  
 Pro Arg Phe Leu Thr Phe Leu Gly Ser Pro Arg Leu Arg Ala Leu  
 65 70 75  
 Ala Arg Gly Leu Ser Pro Ala Tyr Leu Arg Phe Gly Gly Thr Lys  
 80 85 90  
 Thr Asp Phe Leu Ile Phe Asp Pro Asp Lys Glu Pro Thr Ser Glu  
 95 100 105  
 Glu Arg Ser Tyr Trp Lys Ser Gln Val Asn His Asp Ile Cys Arg  
 110 115 120  
 Ser Glu Pro Val Ser Ala Ala Val Leu Arg Lys Leu Gln Val Glu  
 125 130 135  
 Trp Pro Phe Gln Glu Leu Leu Leu Arg Glu Gln Tyr Gln Lys  
 140 145 150

|                                                             |     |     |     |
|-------------------------------------------------------------|-----|-----|-----|
| Glu Phe Lys Asn Ser Thr Tyr Ser Arg Ser Ser Val Asp Met Leu | 155 | 160 | 165 |
| Tyr Ser Phe Ala Lys Cys Ser Gly Leu Asp Leu Ile Phe Gly Leu | 170 | 175 | 180 |
| Asn Ala Leu Leu Arg Thr Pro Asp Leu Arg Trp Asn Ser Ser Asn | 185 | 190 | 195 |
| Ala Gln Leu Leu Leu Asp Tyr Cys Ser Ser Lys Gly Tyr Asn Ile | 200 | 205 | 210 |
| Ser Trp Glu Leu Gly Asn Glu Pro Asn Ser Phe Trp Lys Lys Ala | 215 | 220 | 225 |
| His Ile Leu Ile Asp Gly Leu Gln Leu Gly Glu Asp Phe Val Glu | 230 | 235 | 240 |
| Leu His Lys Leu Leu Gln Arg Ser Ala Phe Gln Asn Ala Lys Leu | 245 | 250 | 255 |
| Tyr Gly Pro Asp Ile Gly Gln Pro Arg Gly Lys Thr Val Lys Leu | 260 | 265 | 270 |
| Leu Arg Ser Phe Leu Lys Ala Gly Gly Glu Val Ile Asp Ser Leu | 275 | 280 | 285 |
| Thr Trp His His Tyr Tyr Leu Asn Gly Arg Ile Ala Thr Lys Glu | 290 | 295 | 300 |
| Asp Phe Leu Ser Ser Asp Ala Leu Asp Thr Phe Ile Leu Ser Val | 305 | 310 | 315 |
| Gln Lys Ile Leu Lys Val Thr Lys Glu Ile Thr Pro Gly Lys Lys | 320 | 325 | 330 |
| Val Trp Leu Gly Glu Thr Ser Ser Ala Tyr Gly Gly Gly Ala Pro | 335 | 340 | 345 |
| Leu Leu Ser Asn Thr Phe Ala Ala Gly Phe Met Trp Leu Asp Lys | 350 | 355 | 360 |
| Leu Gly Leu Ser Ala Gln Met Gly Ile Glu Val Val Met Arg Gln | 365 | 370 | 375 |
| Val Phe Phe Gly Ala Gly Asn Tyr His Leu Val Asp Glu Asn Phe | 380 | 385 | 390 |
| Glu Pro Leu Pro Asp Tyr Trp Leu Ser Leu Leu Phe Lys Lys Leu | 395 | 400 | 405 |
| Val Gly Pro Arg Val Leu Leu Ser Arg Val Lys Gly Pro Asp Arg | 410 | 415 | 420 |
| Ser Lys Leu Arg Val Tyr Leu His Cys Thr Asn Val Tyr His Pro | 425 | 430 | 435 |
| Arg Tyr Gln Glu Gly Asp Leu Thr Leu Tyr Val Leu Asn Leu His | 440 | 445 | 450 |
| Asn Val Thr Lys His Leu Lys Val Pro Pro Pro Leu Phe Arg Lys | 455 | 460 | 465 |
| Pro Val Asp Thr Tyr Leu Leu Lys Pro Ser Gly Pro Asp Gly Leu | 470 | 475 | 480 |
| Leu Ser Lys Ser Val Gln Leu Asn Gly Gln Ile Leu Lys Met Val | 485 | 490 | 495 |
| Asp Glu Gln Thr Leu Pro Ala Leu Thr Glu Lys Pro Leu Pro Ala | 500 | 505 | 510 |
| Gly Ser Ala Leu Ser Leu Pro Ala Phe Ser Tyr Gly Phe Phe Val | 515 | 520 | 525 |
| Ile Arg Asn Ala Lys Ile Ala Ala Cys Ile                     | 530 | 535 |     |

## (2) INFORMATION FOR SEQ ID NO:45:

## (i) SEQUENCE CHARACTERISTICS:

|     |         |              |
|-----|---------|--------------|
| (A) | LENGTH: | 2396         |
| (B) | TYPE:   | nucleic acid |

(C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear  
 (xl) SEQUENCE DESCRIPTION: SEQ ID NO:45

|                                                             |            |     |
|-------------------------------------------------------------|------------|-----|
|                                                             | TT TCT AGT | 8   |
| TGC TTT TAG CCA ATG TCG GAT CAG GTT TTT CAA GCG ACA AAG AGA |            | 53  |
| TAC TGA GAT CCT GGG CAG AGG ACA TCC TAG CTC GGT CAG ATT TGG |            | 98  |
| GCA GGC TCA AGT GAC CAG TGT CTT AAG GCA GAA GGG AGT CCG GGT |            | 143 |
| AGG GTC TGG CTG AAC CCT CAA CCG GGG CTT TTA ACT CAG GGT CTA |            | 188 |
| GTC CTG GCG CCA AAT GGA TGG GAC CTA GAA AAG GTG ACA GAG TGC |            | 233 |
| GCA GGA CAC CAG GAA GCT GGT CCC ACC CCT GCG CGG CTC CCG GGC |            | 278 |
| GCT CCC TCC CCA GGC CTC CGA GGA TCT TGG ATT CTG GCC ACC TCC |            | 323 |
| GCA CCC TTT GGA TGG GTG TGG ATG ATT TCA AAA GTG GAC GTG ACC |            | 368 |
| GCG GCG GAG GGG AAA GCC AGC ACG GAA ATG AAA GAG AGC GAG GAG |            | 413 |
| GGG AGG CCG GGG AGG GGA GGG CCG TAG GGA GGG ACT CCC GGG AGG |            | 458 |
| GGT GGG AGG GAT GGA GCG CTG TGG GAG GGT ACT GAG TCC TGG CGC |            | 503 |
| CAG AGG CGA AGC AGG ACC GGT TGC AGG GGG CTT GAG CCA GCG CGC |            | 548 |
| CGG CTG CCC CAG CTC TCC CCG CAG CCG GCG GTC CAG CCA GGT GGG |            | 593 |
| ATG CTG AGG CTG CTG CTG TGG CTC TGG GGG CCG CTC GGT GCC     |            | 638 |
| Met Leu Arg Leu Leu Leu Trp Leu Trp Gly Pro Leu Gly Ala     |            |     |
| 5 10 15                                                     |            |     |
| CTG GCC CAG GGC GCC CCC GCG GGG ACC GCG CCG ACC GAC GAC GTG | 683        |     |
| Leu Ala Gln Gly Ala Pro Ala Gly Thr Ala Pro Thr Asp Asp Val |            |     |
| 20 25 30                                                    |            |     |
| GTA GAC TTG GAG TTT TAC ACC AAG CCG CCG CTC CGA AGC GTG AGT | 728        |     |
| Val Asp Leu Glu Phe Tyr Thr Lys Arg Pro Leu Arg Ser Val Ser |            |     |
| 35 40 45                                                    |            |     |
| CCC TCG TTC CTG TCC ATC ACC ATC GAC GCC AGC CTG GCC ACC GAC | 773        |     |
| Pro Ser Phe Leu Ser Ile Thr Ile Asp Ala Ser Leu Ala Thr Asp |            |     |
| 50 55 60                                                    |            |     |
| CCG GCG TTC CTC ACC TTC CTG GGC TCT CCA AGG CTC CGT GCT CTG | 818        |     |
| Pro Arg Phe Leu Thr Phe Leu Gly Ser Pro Arg Leu Arg Ala Leu |            |     |
| 65 70 75                                                    |            |     |
| GCT AGA GGC TTA TCT CCT GCA TAC TTG AGA TTT GGC GGC ACA AAG | 863        |     |
| Ala Arg Gly Leu Ser Pro Ala Tyr Leu Arg Phe Gly Gly Thr Lys |            |     |
| 80 85 90                                                    |            |     |
| ACT GAC TTC CTT ATT TTT GAT CCG GAC AAG GAA CCG ACT TCC GAA | 908        |     |
| Thr Asp Phe Leu Ile Phe Asp Pro Asp Lys Glu Pro Thr Ser Glu |            |     |
| 95 100 105                                                  |            |     |
| GAA AGA AGT TAC TGG AAA TCT CAA GTC AAC CAT GAT ATT TGC AGG | 953        |     |
| Glu Arg Ser Tyr Trp Lys Ser Gln Val Asn His Asp Ile Cys Arg |            |     |
| 110 115 120                                                 |            |     |
| TCT GAG CCG GTC TCT GCT GCG GTG TTG AGG AAA CTC CAG GTG GAA | 998        |     |
| Ser Glu Pro Val Ser Ala Ala Val Leu Arg Lys Leu Gln Val Glu |            |     |
| 125 130 135                                                 |            |     |
| TGG CCC TTC CAG GAG CTG TTG CTG CTC CGA GAG CAG TAC CAA AAG | 1043       |     |
| Trp Pro Phe Gln Glu Leu Leu Leu Leu Arg Glu Gln Tyr Gln Lys |            |     |
| 140 145 150                                                 |            |     |



GAG TTC AAG AAC AGC ACC TAC TCA AGA AGC TCA GTG GAC ATG CTC 1088  
 Glu Phe Lys Asn Ser Thr Tyr Ser Arg Ser Ser Val Asp Met Leu  
 155 160 165

TAC AGT TTT GCC AAG TGC TCG GGG TTA GAC CTG ATC TTT GGT CTA 1133  
 Tyr Ser Phe Ala Lys Cys Ser Gly Leu Asp Leu Ile Phe Gly Leu  
 170 175 180

AAT GCG TTA CTA CGA ACC CCA GAC TTA CGG TGG AAC AGC TCC AAC 1178  
 Asn Ala Leu Leu Arg Thr Pro Asp Leu Arg Trp Asn Ser Ser Asn  
 185 190 195

GCC CAG CTT CTC CTT GAC TAC TGC TCT TCC AAG GGT TAT AAC ATC 1223  
 Ala Gln Leu Leu Leu Asp Tyr Cys Ser Ser Lys Gly Tyr Asn Ile  
 200 205 210

TCC TGG GAA CTG GGC AAT GAG CCC AAC AGT TTC TGG AAG AAA GCT 1268  
 Ser Trp Glu Leu Gly Asn Glu Pro Asn Ser Phe Trp Lys Lys Ala  
 215 220 225

CAC ATT CTC ATC GAT GGG TTG CAG TTA GGA GAA GAC TTT GTG GAG 1313  
 His Ile Leu Ile Asp Gly Leu Gln Leu Gly Glu Asp Phe Val Glu  
 230 235 240

TTG CAT AAA CTT CTA CAA AGG TCA GCT TTC CAA AAT GCA AAA CTC 1358  
 Leu His Lys Leu Leu Gln Arg Ser Ala Phe Gln Asn Ala Lys Leu  
 245 250 255

TAT GGT CCT GAC ATC GGT CAG CCT CGA GGG AAG ACA GTT AAA CTG 1403  
 Tyr Gly Pro Asp Ile Gly Gln Pro Arg Gly Lys Thr Val Lys Leu  
 260 265 270

CTG AGG AGT TTC CTG AAG GCT GGC GGA GAA GTG ATC GAC TCT CTT 1448  
 Leu Arg Ser Phe Leu Lys Ala Gly Gly Glu Val Ile Asp Ser Leu  
 275 280 285

ACA TGG CAT CAC TAT TAC TTG AAT GGA CGC ATC GCT ACC AAA GAA 1493  
 Thr Trp His His Tyr Tyr Leu Asn Gly Arg Ile Ala Thr Lys Glu  
 290 295 300

GAT TTT CTG AGC TCT GAT GCG CTG GAC ACT TTT ATT CTC TCT GTG 1538  
 Asp Phe Leu Ser Ser Asp Ala Leu Asp Thr Phe Ile Leu Ser Val  
 305 310 315

CAA AAA ATT CTG AAG GTC ACT AAA GAG ATC ACA CCT GGC AAG AAG 1583  
 Gln Lys Ile Leu Lys Val Thr Lys Glu Ile Thr Pro Gly Lys Lys  
 320 325 330

GTC TGG TTG GGA GAG ACG AGC TCA GCT TAC GGT GGC GGT GCA CCC 1628  
 Val Trp Leu Gly Glu Thr Ser Ser Ala Tyr Gly Gly Gly Ala Pro  
 335 340 345

TTG CTG TCC AAC ACC TTT GCA GCT GGC TTT ATG TGG CTG GAT AAA 1673  
 Leu Leu Ser Asn Thr Phe Ala Ala Gly Phe Met Trp Leu Asp Lys  
 350 355 360

|                                                             |      |
|-------------------------------------------------------------|------|
| TTG GGC CTG TCA GCC CAG ATG GGC ATA GAA GTC GTG ATG AGG CAG | 1718 |
| Leu Gly Leu Ser Ala Gln Met Gly Ile Glu Val Val Met Arg Gln |      |
| 365 370 375                                                 |      |
| GTG TTC TTC GGA GCA GGC AAC TAC CAC TTA GTG GAT GAA AAC TTT | 1763 |
| Val Phe Phe Gly Ala Gly Asn Tyr His Leu Val Asp Glu Asn Phe |      |
| 380 385 390                                                 |      |
| GAG CCT TTA CCT GAT TAC TGG CTC TCT CTT CTG TTC AAG AAA CTG | 1808 |
| Glu Pro Leu Pro Asp Tyr Trp Leu Ser Leu Leu Phe Lys Lys Leu |      |
| 395 400 405                                                 |      |
| GTA GGT CCC AGG GTG TTA CTG TCA AGA GTG AAA GGC CCA GAC AGG | 1853 |
| Val Gly Pro Arg Val Leu Leu Ser Arg Val Lys Gly Pro Asp Arg |      |
| 410 415 420                                                 |      |
| AGC AAA CTC CGA GTG TAT CTC CAC TGC ACT AAC GTC TAT CAC CCA | 1898 |
| Ser Lys Leu Arg Val Tyr Leu His Cys Thr Asn Val Tyr His Pro |      |
| 425 430 435                                                 |      |
| CGA TAT CAG GAA GGA GAT CTA ACT CTG TAT GTC CTG AAC CTC CAT | 1943 |
| Arg Tyr Gln Glu Gly Asp Leu Thr Leu Tyr Val Leu Asn Leu His |      |
| 440 445 450                                                 |      |
| AAT GTC ACC AAG CAC TTG AAG GTA CCG CCT CCG TTG TTC AGG AAA | 1988 |
| Asn Val Thr Lys His Leu Lys Val Pro Pro Pro Leu Phe Arg Lys |      |
| 455 460 465                                                 |      |
| CCA GTG GAT ACG TAC CTT CTG AAG CCT TCG GGG CGG GAT GGA TTA | 2033 |
| Pro Val Asp Thr Tyr Leu Leu Lys Pro Ser Gly Pro Asp Gly Leu |      |
| 470 475 480                                                 |      |
| CTT TCC AAA TCT GTC CAA CTG AAC GGT CAA ATT CTG AAG ATG GTG | 2078 |
| Leu Ser Lys Ser Val Gln Leu Asn Gly Gln Ile Leu Lys Met Val |      |
| 485 490 495                                                 |      |
| GAT GAG CAG ACC CTG CCA GCT TTG ACA GAA AAA CCT CTC CCC GCA | 2123 |
| Asp Glu Gln Thr Leu Pro Ala Leu Thr Glu Lys Pro Leu Pro Ala |      |
| 500 505 510                                                 |      |
| GGA AGT GCA CTA AGC CTG CCT GCC TTT TCC TAT GGT TTT TTT GTC | 2168 |
| Gly Ser Ala Leu Ser Leu Pro Ala Phe Ser Tyr Gly Phe Phe Val |      |
| 515 520 525                                                 |      |
| ATA AGA AAT GCC AAA ATC GCT GCT TGT ATA TGA AAA TAA AAG GCA | 2213 |
| Ile Arg Asn Ala Lys Ile Ala Ala Cys Ile                     |      |
| 530 535                                                     |      |
| TAC GGT ACC CCT GAG ACA AAA GCC GAG GGG GGT GTT ATT CAT AAA | 2258 |
| ACA AAA CCC TAG TTT AGG AGG CCA CCT CCT TGC CGA GTT CCA GAG | 2303 |
| CTT CGG GAG GGT GGG GTA CAC TTC AGT ATT ACA TTC AGT GTG GTG | 2348 |
| TTC TCT CTA AGA AGA ATA CTG CAG GTG GTG ACA GTT AAT AGC ACT | 2393 |
| CTG                                                         | 2396 |

## (2) INFORMATION FOR SEQ ID NO:46:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 385

(B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46  
 CGGCCGCTGC TGCTGCTGTG GCTCTGGGG CGGCTCCGTG CCCTGACCCA 50  
 AGGCACCTCG GCGGGGACCG CGCCGACCAA AGACGTGGTG GACTTGGAGT 100  
 TTACACCAA GAGGCTATTC CAAAGCGTGA GTCCCTCGTT CCTGTCCATC 150  
 ACCATCGACG CCAGTCTGGC CACCGACCTT CGGTTCCTCA CCTTCTGTAG 200  
 CTCTCCACGG CTTCGAGGCC TGTCTAGAGG CTTATCTCCT GGTACTTGA 250  
 GATTTGGCGG CACCAAGACT GACTTCCTTA TTTTGTATCC CAACAACGAA 300  
 CCCACCTCTG AAGAAAGAAG TTAAGTGCAT TCTCAAGACA ACAATGATAT 350  
 TTGCGGGTCT GACCGGGTCT CCGCTGACGT GTTGA 385

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 541  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47  
 AAATCAGGAC ATATCTTCA CTTATTGCC TCTTGGTCAT ATTGGAGGCA 50  
 TTTGTATICA TTTTAAATAA CCTCAAAAT AGTGCAITGA AAGTGCTAAG 100  
 CGTCATTTCG CACATGCTGC CATTAACCTG CACCACTGCG AGTGGTCTAC 150  
 TTAGAGAACA CCGCACTGGA TGTAAACACT GAAGCGCGTG CCCCGCCCTC 200  
 CCGAGGCTCT GGATCCAGCG TTGAAGCTTG CCGCGCCCTC CCGAGGCTCT 250  
 GGATCCAGCA CTGGAGCATG CCCCGCCCTC CCGAGGCTCT GGAGCTTGCT 300  
 AAGGAGTCCG CTCCTTACCG CTGGGGTTTT GCTTATTCT TATGAATGAC 350  
 ACCCCTGACC GCTTCTGCTC CAGGGGTACT GTAATGCCTT TTATTTTCAI 400  
 ATACAAGCTG CGATTTTGGC ATTCTTATG ACAAAAAACC CATAGGAAAA 450  
 GCGGGGCACG CTTAGTGAGC TTCTCGCGGG GAGAGGTTTT TCTGTTAGAG 500  
 CTGGCANGGT CTGCTCATCG ACCATCTTCA GGCCTCGTGC C 541